Genetic and environmental predictors of psychiatric disorders and related traits

Power, Robert

Awarding institution:
King's College London

Genetic and environmental predictors of psychiatric disorders and related traits

Robert Power

MRC Social, Genetic & Developmental Psychiatry Centre

Institute of Psychiatry

King's College London

Submitted for the degree of Doctor of Philosophy

Acknowledgements

Many people have contributed towards this work, providing guidance and support. First and foremost were my supervisors, Cathryn Lewis and Peter McGuffin. When I began my PhD I was told by many staff and students at the SGDP how lucky I was to have such amazing supervisors. Little did I understand then, but they proved it true a hundred times over these last 4 years. I'd like to thank them for the excellent supervision, support and mentoring they've provided. They have allowed me to follow my own interests and projects, giving advice not just on scientific training but also on how to conduct oneself as an ethical and successful researcher.

I'd also like to thank the numerous collaborators who have contributed to the papers and projects of my PhD. At the SGDP special thanks goes to Rudolf Uher, Gerome Breen and Robert Plomin. Externally, to Nick Martin and Naomi Wray in Brisbane and Paul Lichtenstein at the Karolinska Institutet, whose invitations to visit their institutes made the work in this thesis possible. Also thanks to the Medical Research Council for funding my PhD and the SGDP as a whole.

Lastly, I'd like to thank all the support I've had from my friends and family. Many amazing memories have been made with those of you at the SGDP, and I can definitely say my friends here made my degree. I'd name you all individually, but it'd end up being the whole department and I'll run out of adjectives half way through. Thanks to my father for inspiring my interest in psychiatry and science from an early age, my mother for always kicking me into gear when needed (only sometimes literally), my brother for being the best housemate possible while undergoing his own valiant attempts to become a human mechanic/"real doctor", and…. well nothing to my little sister. Georgie, I will never respect you.

In Egypt's sandy silence all alone

Still searching how best to atone

Sweet Bubastis so loyal

Only for you did I toil

On a thesis of stone that time will foil

Contents

Abstract

Psychiatric disorders pose a major economic and health burden worldwide. Family and twin studies indicate strong genetic influences, with estimates suggesting substantial variance in liability is heritable. However known genetic risk variants for psychiatric disorders explain only a small fraction of the heritability estimated from twin studies, leaving the underlying genetic aetiology largely unknown. Given the wide range in prevalence, age at onsets, and gender ratios seen across psychiatric disorders it is reasonable to expect that different genetic architectures exist for each disorder. Thus the underlying genetic architecture of psychiatric disorders remains an open question, with large potential implications for identifying predictive genetic risk variants, redefining diagnostic criteria, and developing novel drug targets.

This thesis focuses on combining epidemiological, genetic and environmental data to answer the questions surrounding the genetic architectures of psychiatric disorders. Initial analysis of epidemiological measures surprisingly identified that major depression was not under negative selection, a finding that was then confirmed through analysis of genotypic data. These results suggested the potential role of gene-environment interactions as an adaptive mechanism for variants contributing risk to depression, and were followed up by studies focusing on identifying such interactions. The conclusion of the thesis was though no gene-environmental interaction could be found to explain how the risk variants for major depression avoided negative selection, this was likely due in part to substantial gene-environment correlation in the reporting and experiencing of environmental risk factors in psychiatric disorders which would confound such analyses. Further these gene-environment correlations likely confound epidemiological associations identifying 'environmental' risks, such as between cannabis and schizophrenia.

Statement of authorship

All work in this thesis was performed and written by Robert Power, with the following exceptions:

1) The published papers were both circulated amongst co-authors and underwent peer review prior to publication, leading to editing of manuscript and suggestions for additional analyses.
2) The datasets used throughout the published papers were collected by others prior to analysis and underwent some quality control and cleaning (true for both phenotypic, epidemiological and genetic data).

1. Introduction

1.1.1 Outline

Psychiatric disorders pose a major economic and public health burden in developing and developed countries alike [Vos and others 2012]. Disorders such as schizophrenia, bipolar disorder, anorexia, autism and major depression are highly prevalent, with early mean ages of onset that lead to high levels of disability and morbidity, reduced reproductive fitness, and mortality. Family and twin studies indicate strong genetic influences, with estimates suggesting substantial variance in liability is heritable [e.g. Lee and others 2013b; Lichtenstein and others 2009]. This combination has long been noted as a paradox since natural selection should deplete genetic variation associated with reduced reproductive fitness [Keller and Miller 2006; Uher 2009]. This is further complicated as known genetic risk variants for psychiatric disorders explain only a small fraction of the heritability estimated from twin studies, leaving the underlying genetic aetiology largely unknown [Sullivan and others 2012]. Thus the underlying genetic architecture of psychiatric disorders remains an open question in the field, with large potential implications for identifying predictive genetic risk variants, redefining diagnostic criteria, and developing novel drug targets. This thesis will begin by outlining the nine major neuro-psychiatric disorders, then describing the relevant epidemiological and molecular genetic findings for each. The introduction will finish with summary of the publications which will compose the thesis.

1.1.2 Summary of neuropsychiatric disorders

1.1.2.1 Schizophrenia

Schizophrenia is a psychotic disorder that affects around 1% of the population [McGrath and others 2008], with a slightly higher prevalence in males than females. It is characterised by three main clusters of symptoms: positive symptoms such as hallucinations and paranoia; negative symptoms of anhedonia and catatonia; and disorganised thought processes with language and cognitive impairment [APA 1994]. Onset occurs most frequently in adolescents or early adulthood, on

average slightly earlier for males [Faraone and others 1994]. It has been associated with urbanicity, migrant status, latitude, and season of birth [McGrath and others 2008]. Cannabis use has also been associated with a twofold increased risk of schizophrenia [Arseneault and others 2004]. Schizophrenia is associated with increased mortality [Joukamaa and others 2001], in part due the suicide related symptoms of the disorder.

1.1.2.2 Bipolar disorder

Bipolar disorder is a mood disorder, similar to schizophrenia in that it often includes psychosis and affects around 1% of the population [Bebbington and Ramana 1995]. However the positive and negative symptoms of bipolar disorder manifest in a more cyclical manner, and it lacks the same level of cognitive impairment. Individuals switch between episodes of manic behaviour to episodes of depression [APA 1994]. Two subtypes are often defined: bipolar disorder type 1 where there are clear manic and depressive episodes; and bipolar disorder type 2 where less clear hypomania episodes exist among periods of depression. Bipolar disorder has been associated with risk factors such as divorce, financial problems and job loss [Hosang and others 2012], though this may be driven by prodromal symptoms as these are all common side effects of the reckless sexual and financial behaviour that can accompany manic episodes. Unlike schizophrenia it does not show a robust association with lower socioeconomic status, with some studies even finding the opposite direction of affect [reviewed by Bebbington and Ramana 1995]. Interestingly, some studies have suggested an association with increased cognitive abilities in individuals with bipolar disorder [Gale and others 2013; MacCabe and others 2010]. Despite this, increased mortality exists and suicide is prevalent among those affected [Osby and others 2001].

1.1.2.3 Major depressive disorder

Major depressive disorder (MDD), also known as unipolar depression, is the most highly prevalent psychiatric disorder (alongside the often co-morbid anxiety disorders) affecting ~15% of individuals over their lifetime usually at twice the prevalence in women as in men [Hasin and

others 2005; Kessler and others 2003]. With most individuals experiencing recurrent episodes throughout life [Mueller and others 1999], it is now the 2nd leading cause of disability worldwide [Vos and others 2012]. It is a highly heterogeneous disorder, defined by reduced mood and energy, inability to experience enjoyment, changes to diet and sleep patterns, feelings of guilt or worthlessness, and suicidal thoughts [APA 1994]. Along with excess mortality and increased risk of suicide [Angst and others 2002], it is associated with worse clinical outcomes when co-morbid with health problems such as cardiovascular disease and cancer [Barth and others 2004; Satin and others 2009]. MDD can occur at any age with many life events associated with onset such as puberty, menopause, childhood maltreatment, childbirth, and divorce [Angold and others 1998; Freeman and others 2004; Greenwald and others 1989; Hosang and others 2012; Thapar and others 2012a].

1.1.2.4 Autism

Autism is a lifelong developmental disorder that is manifests in early childhood, with a prevalence of ~0.1% and found around 4 times as frequently in males than females [Fombonne 2005]. Prevalence has increased over recent decades though this is believed to be driven largely by improved ascertainment and broader diagnostic categories [Rutter 2005]. Symptoms fall into three broad categories: social impairment; reduced ability to communicate; and repetitive behaviour [APA 1994]. Autism is known to exist on a spectrum of severity and affected individuals will not necessarily have symptoms from all three categories, though severe impairment in one will of course impair the others. It is also often co-morbid with intellectual disability.

1.1.2.5 Attention deficit hyperactivity disorder

Attention-deficit hyperactivity disorder (ADHD) usually manifests in childhood, and is defined by developmentally inappropriate levels of hyperactivity, impulsiveness and inattention [APA 1994]. These symptoms are generally considered to give rise to three subtypes of ADHD: predominantly inattentive type (high inattention, low hyperactivity), predominantly hyperactive-impulsive type

(low inattention, high hyperactivity), and combined type (high on both symptom dimensions). ADHD was initially believed to be a childhood disorder but is now recognised as frequently spanning from preschool age into adolescence and adulthood [Barkley 2008]. The prevalence of ADHD has been estimated at around ~5% in children and adolescents [Polanczyk and others 2007] and at around 2.5% in the adult population [Simon and others 2009]. It has been suggested that ADHD is associated with lower cognitive ability [Kuntsi and others 2004]. Associations with prenatal environment, such as maternal alcohol use or stress, have been reported but are not firmly established [Thapar and others 2012b].

1.1.2.6 Eating disorders

Anorexia nervosa and bulimia nervosa represent the two most prominent eating disorders. Anorexia nervosa is characterised by extreme dieting and avoidance of food while bulimia nervosa is characterised by binge eating followed by purging, with a combined prevalence of around 1% and anorexia nervosa being the rarer of the two [Hoek and van Hoeken 2003; Hudson and others 2007]. Both have a much higher prevalence in women, usually with onset in the late teens [Lucas and others 1991]. Prevalence in men may be higher than estimated though due to underreporting [Strother and others 2012], and it is known that prevalence of anorexia varies greatly across countries and cultures [Hoek 2006; Pavlova and others 2009]. Anorexia is a particularly serious disorder due to its extremely high mortality rate, the highest of any psychiatric disorder [Sullivan 1995; Zipfel and others 2000].

1.1.2.7 Substance abuse and addiction

Substance abuse covers a large range of disorders that each vary greatly in their prevalence, usually as a result of the availability of the substance both in terms of geography and generation. The most common substances are those legally available such as alcohol and tobacco, with lifetime prevalences estimated at 17% and 24% respectively [Breslau and others 2001; Hasin and others 2007], but extend to illicit drugs such as cannabis, opiates, or cocaine. Addiction is a complex trait

that can be conceived on many levels that may not apply equally to every substance, such as the willingness to try new or illegal experiences, choice of social group and peers, ability to tolerate the substance, susceptibility to addiction in terms of personality, and susceptibility to addiction based on biological disposition to the substance. One of the more consistent findings from studies of addition and substance abuse is that males are at greater risk [Becker and Hu 2008].

1.1.2.8 Alzheimer's disease

Alzheimer's disease is the most common form of dementia, accounting for up to 75% of cases [Lobo and others 2000]. It is characterised by rapid cognitive decline with an onset in late life, though early-onset forms exist. Symptoms include confusion, mood swings, language problems and memory loss. It is associated with a wide variety of risk factors [reviewed by Qiu and others 2009], including alcohol use, blood pressure, and poor social or mental activity throughout life. Ultimately cognitive functioning decreases until death.

1.1.3 Expectations

Clearly psychiatric disorders represent a wide variety of behavioural problems and impairments. They range greatly in prevalence from 0.1% to 15%, often appear much more frequently in women (e.g. anorexia and depression) or men (autism and substance abuse), and have onsets that range from just after birth to old age. Given the wide range in prevalence, age at onsets, and gender ratios seen across psychiatric disorders it is reasonable to expect that different genetic architectures exist for each disorder. In the next two sections I will evaluate the epidemiological and molecular evidence for the underlying genetics of these disorders, and build a consensus on where the causal variants exist.

1.2 Genetic epidemiology

Epidemiological studies can provide insight into the underlying genetics of any trait without requiring molecular genetic data. Twin and adoption studies can be used to test the relative impact of genes and environment on a trait, estimating the heritable component to variation between

individuals. Impact on fecundity gives us an insight into the selection on a trait, and so how likely it is that causal variants can survive in the population. Paternal aging is associated with increased burden of *de novo* mutations in the sperm, meaning an association with older fathers may reflect the importance of novel mutations. And lastly, non-random mating in the population with respect to a trait gives us insight into the likely distribution of causal variants. In this section I will expand on each of these epidemiological features and their bearing on psychiatric disorders.

1.2.1 Twin studies

The format of a twin study is intuitively appealing, as it naturally controls several variables. Twins are born at the same time in the same family, and hence should share a similar environment. They also share genes, 100% in identical or monozygotic twins (MZ) and 50% in terms of fraternal or dizygotic twins (DZ). These percentages refer of course to genetic differences relative to the general population (where one is 0% identical to an unrelated individual), as most of the human genome is unchanging across all individuals. This difference in genetic similarity between twins is due to MZ twins coming from a single fertilised egg splitting into two, while DZ come from two separately but simultaneously fertilised eggs. A comparison of concordance rates between and within DZ and MZ twin pairs gives an insight into the role of genetics in the disease. If a disease was entirely genetic in nature, it would be expected that MZ twins were 100% concordant, and DZ twins 50% discordant reflecting their genetic differences. Even when that's not the case, the level of genetic difference between MZ and DZ twins can be used to estimate genetic, shared and non-shared environmental factors through simple simultaneous equations when correlations of twins is known. This follows from the equations below.

$rMZ = h^2 + c^2$

$rDZ = h^2/2 + c^2$

$Hence\ 2(rMZ - rDZ) = h^2$

Here the rMZ reflects the correlation between MZ twins, which should be the heritability of the trait (h^2) plus any shared environmental influences (c^2). rDZ is the correlation between DZ twins which is the same as that of MZ, except reflecting only half the heritability as they share only half their genetic predisposition. As such, the correlations between MZ and DZ twins can be used to determine the heritability. The use of more advanced model fitting can then expand on these calculations, and provide confidence intervals for the estimates of genes and environment on the trait.

The majority of psychiatric traits have high heritabilities, though range from 0.37-0.81. The highest heritabilities are in schizophrenia, bipolar disorder, ADHD and autism. In reviews of 12 twin studies of schizophrenia and bipolar disorders the heritabilities were estimated to be 0.81 and 0.75 respectively [Smoller and Finn 2003; Sullivan and others 2003], though population studies give lower estimates at around 0.6 as well as reporting significant genetic overlap between them [Lichtenstein and others 2009; Wray and Gottesman 2012]. A meta-analysis of 20 studies of ADHD showed a heritability of 0.76 [Faraone and others 2005], though estimates have been found to be higher in children than in adults [Boomsma and others 2010; Larsson and others 2013]. A similar meta-analysis of 30 twin studies of autism estimated a heritability of 0.80 [Ronald and Hoekstra 2011], with some evidence for substantial overlap in the heritability for severe autism cases as those on a more moderate spectrum [Robinson and others 2011]. Studies of Alzheimer's disease and anorexia nervosa have suggested moderate heritabilities, of 0.58 and 0.56 respectively [Bulik and others 2007; Gatz and others 2006]. The heritability of substance abuse shows large ranges, perhaps due to the differing definitions of addiction across cultures and countries, or the availability of addictive substances. Alcohol addiction ranges from 0.48–0.66, nicotine from 33–71%, cannabis 51–59%, and cocaine from 42 to 79% [reviewed by Agrawal and Lynskey 2006]. The lowest heritability are found for major depression at around 0.35 in both twin and population

studies [Sullivan and others 2000; Wray and Gottesman 2012], though some studies suggest it is substantially higher in early-onset recurrent cases [Levinson 2006; McGuffin and others 1996].

Note that these estimates reflect additive genetic effects, and so ignore gene-environment or gene-gene interactions, dominant and recessive effects, and *de novo* mutations. They also rely on several assumptions, most importantly the equal environments assumption that MZ twins are not treated more similarly than DZ twins. Twin estimates also assume no assortative mating on the trait studies and that twins accurately represent this trait in the population, both of which may vary across disorders. Hence, twin estimates of heritability are not without some caveats when used as a benchmark for the role of genetics in any given disorder.

1.2.2 Reproductive fitness

Natural selection shapes the genetic architecture of any heritable trait. A genetic variant's ability to stay at a given frequency is dependent on its impact on an organism's reproductive fitness. The mutation-selection balance hypothesis suggests that deleterious genetic variants remain in the gene pool as new mutations are introduced at a pace equal to the selection against existing copies of the variant. An extension of this is the common disease common variant (CDCV) hypothesis where by a large number of risk variants with individually very small effects could escape selection. Antagonistic pleiotropy may also play a role, whereby genes that increase liability to a deleterious trait are beneficial under some circumstances, compensating for the negative selection in affected individuals. Selection shapes the genomic landscape of any trait, such as the number of risk alleles, their effect size, and their penetrance.

Calculating the impact of psychiatric disorders on reproductive fitness is usually performed presenting a fecundity ratio (FR), comparing the fecundity of the disorder to that of the general population. Studies have largely focused on schizophrenia, and their results were meta-analysed in a recent review [Bundy and others 2011]. Patients with schizophrenia were found to have

significantly reduced fitness (FR=0.39) while siblings showed a slight reduction (FR=0.96), with the effect being larger in males in both instances. They reported that parents of schizophrenic individuals showed no significant difference in FR, while one of the included studies reported that male offspring of schizophrenic individuals show reduced fecundity [Svensson and others 2007]. Outside of schizophrenia, bipolar disorder has not as consistently shown strong associations with reduced fecundity [MacCabe and others 2009; Mansour and others 2011; McGrath and others 1999], though there is some evidence for reduced fecundity in autism, eating disorders and addictions [Uher 2009].

Some studies have focused on potential benefits of psychiatric risk variants, specifically creativity. Several early attempts were made to quantify the overlap of creativity and mental health, looking for an increased levels of creative professions in those affected by and the relatives of those with psychiatric disorders [Andreasen 1987; Herbert 1959; Jamison 1989; Juda 1949; Ludwig 1992; Post 1994]. For the most part these studies reported such an association, though were restricted by biased ascertainment and small sample sizes. In a more recent family study of 300,000 individuals [Kyaga and others 2011], those with bipolar disorder and healthy siblings of people with schizophrenia or bipolar disorder were overrepresented in creative professions, but not for individuals with or related to those with major depression. In a subsequent larger study, Kyaga and others [2013] found that, except for bipolar disorder, individuals in creative professions in general were no more likely to be diagnosed with psychiatric disorders than controls, though being an author was specifically associated with risk of psychiatric illness. Creative professions were found to be more frequent among first-degree relatives of patients with schizophrenia, bipolar disorder, anorexia nervosa, and for siblings of those with autism. These findings have suggested a potentially shared heritability between creativity and psychiatric disorders, particularly bipolar disorder and schizophrenia, might exist.

1.2.3 Paternal age

Paternal age has been associated with increased *de novo* mutations in a father's offspring [Kong and others]. This is assumed to be due to the increased number of DNA replications that occur within sperm precursor cells compared to those that develop into eggs. This provides more opportunities for mutations to occur in male germ-line cells than those in females as they age. Thus an association with paternal age can provide insight into the role of *de novo* mutations in the genetic architecture of a trait. A further mechanism has been suggested, whereby those mutations that benefit sperm in their competition with each other to be passed on to the next generation are not necessarily those that benefit the offspring, and so may lead to deleterious traits [Goriely and others 2013]. This of course could maintain *de novo* mutations across generations.

Early studies showed clear associations between both autism and schizophrenia and paternal age, specifically males fathering offspring over ~40, suggesting *de novo* mutations play a role in these disorders [Brown and others 2002; Durkin and others 2008; Hultman and others 2011; Malaspina and others 2001; Reichenberg and others 2006; Zammit and others 2003]. Bipolar disorder showed a weaker association, though it was stronger in early onset cases [Frans and others 2008]. Follow up studies showed that paternal age acts as a risk factor across generations, with grandpaternal age associated with autism even when paternal age is matched for [Frans and others 2013]. Maternal grandfather's age also appears to contribute to risk in schizophrenia [Frans and others 2011], perhaps suggesting the role of mutations on the X chromosome. Two new studies have examined the impact of paternal age across a variety of disorders in Swedish and Danish populations [D'Onofrio and others 2014; McGrath and others 2014]. These showed paternal age effects across all disorders except eating disorders, including autism, ADHD, bipolar disorder, schizophrenia, suicide attempt, substance abuse, intellectual disability, and low educational attainment. Offspring of very young mothers and fathers were also found to be at increased risk, presumably due to environmental causes rather than genetic differences. It is also possible that older or younger

parents may differ genetically in some way that predisposes their offspring to mental illness. For example, males that take longer to find partners and have children may be further along the schizophrenia spectrum. Thus some have argued that it is actually the father's age at the birth of their first child rather than age at the birth of any individual child that is associated with risk [e.g. Pedersen and others 2014].

1.2.4 Non-random mating

Non-random mating can affect the genetic architecture of a trait in several ways. As mentioned above twin studies assume mating at random, as positive assortative mating (where individuals mate more often with similar individuals) can decrease estimates of heritability by increasing the genetic similarity of DZ twins above 0.5. Positive assortative mating also increases the number of homozygous individuals, pushing individuals to the extreme of the spectrum and potentially impacting the effects of selection. Positive assortative mating has long been shown with measures of personality [Zietsch and others 2011], though it remains unclear to what extent this results from cohabitation of partners. Increased cohabitation and assortative mating has been shown for severe psychiatric disorders, with very high levels reported for schizophrenia and bipolar disorder [Lichtenstein and others 2006; Thomsen and others 2013].

Positive assortative mating can lead to inbreeding defined as when individuals have a greater level of homozygosity in their genome than expected by chance, usually with reference to being identical by descent from a common ancestor. Inbreeding results in homozygosity across the genome, which can expose rare recessive variants that otherwise would not be expressed. Deleterious variants will be selected out of the population but, as the strength of negative and positive selection is greater on those of dominant effect, existing deleterious variants are more likely to be recessive. Hence inbreeding can expose deleterious mutations that otherwise would have been heterozygous. Inbreeding is associated with increased rates of illness, most commonly with reduced cognitive functioning [Jensen 1983; Morton 1978; Woodley 2009].

1.3 Molecular genetic findings

Genetic studies of psychiatric disorders have rapidly expanded over the previous decades, mainly due to the huge changes in technology and cost that have revolutionised the field of genetics as a whole. This ranges from the first linkage studies of the co-segregation of disease and genotype in families, to genome-wide association studies comparing cases and controls for differences at hundreds of thousands or even millions of markers across the genome, to the more recent advances in whole genome sequencing. In this section I will give a brief description of each of these methods and their current findings in psychiatry.

1.3.1 Linkage studies

The first insights into the genetics of psychiatric disorders came from linkage studies. These focus on genotyping genetic polymorphisms within family pedigrees with multiple affected individuals. These studies look for the co-segregation of disease and genotypes across generations of family members, isolating the region in which the causal variant exists. Within this framework the results of multiple pedigrees can be combined to increase power, though this relies on the same mutation having an effect across families. These approaches had huge success with those disorders showing the classic Mendelian inheritance patterns of recessive or dominant effects, of which many include psychiatric aspects such as intellectual disability and autism [Betancur 2011]. For more complex manifestations of psychiatric disorders the results for linkage analyses were less impressive. Large scale meta-analyses of linkage studies for schizophrenia and bipolar disorder showed only one significant finding on chromosomes 2p12-q22.1 predisposing to schizophrenia [Lewis and others 2003; Segurado and others 2003], later found again in a larger meta-analysis including some of the same studies [Ng and others 2009], and was unlikely to explain risk to a large proportion of cases. In contrast to this, Alzheimer's disease showed clear linkage with regions surrounding rare autosomal dominant variants of unusually large effect in the amyloid beta precursor protein (*APP*), presenilin 1 (*PSEN1*) and presenilin 2 (*PSEN2*) genes in families with early success in the severe

form of early-onset Alzheimer's disease [Bertram and Tanzi 2008]. Variants in the apolipoprotein E (*APOE*) were also identified that increased risk in late-onset Alzheimer's disease [Corder and others 1993], a gene that would reappear in genome-wide association studies along with several other common variants that contribute risk [Lambert and others 2013].

1.3.2 Genome-wide association studies (GWAS)

1.3.2.1 Methodology

With the somewhat disappointing results for complex disorders from linkage studies, alternate methods were explored that moved away from the family design best suited for Mendelian disorders. The common disease common variant (CDCV) hypothesis was put forward, broadly suggesting that many high frequency but low effect variants might contribute to risk to complex disorders, thus explaining how they avoided negative selection, manifested in complex inheritance patterns, and allowed for genetic and phenotypic heterogeneity within disorders. To attempt to identify some of these common variants, genome-wide association studies (GWAS) were implemented. These restricted to looking at previously identified common variants (frequency>1%) across the genome, usually in the form of bi-allelic single nucleotide polymorphisms (SNPs). In the case of diseases, DNA samples from case and control groups are analysed using specially developed genotyping chips that contain anywhere from hundreds of thousands to millions of SNPs from across the genome. Each SNP is then tested to see if one variant, or allele, is significantly more frequent among cases than controls. Due to the co-inheritance of segments of the genome over generations, correlations between SNPs allow one to 'tag' local variation in the region. This means that while an association might be found with a SNP, it may actually be tagging a truly causal variant in the region rather than having a direct biological impact on risk. As such only relatively few SNPs need to be genotyped in order to capture a large amount of the common genetic variation across the genome. Further, these correlations between SNPs, or

linkage disequilibrium, also allows for the imputation of additional SNPs not directly genotyped themselves but able to be filled in based on known correlations with those SNPs that have been.

GWAS provided a novel method for identifying risk variants across the genome, without relying on prior biological knowledge of gene function needed to select a gene for candidate studies. This also had the perhaps unexpected outcome of identifying numerous variants in 'gene deserts', without any clear functional impact on gene expression or structure suggesting at a much more complex system. The other benefit was that GWAS could move away from the pedigrees required for linkage analyses, removing some the constraints on recruitment and also alleviating concerns that the ascertainment of familial cases might give insight into a different aetiology compared to sporadic cases. The weakness of the GWAS method was the risk of false positives, the cause being twofold. First, the wide recruitment of unrelated cases and controls runs the risk of subtle population differences between the two groups e.g. an excess of cases from at risk minority groups. This can lead to the identification of SNPs that while highly significant are informative only of ancestry, not the underlying biology of the disease. To account for this many studies restrict recruitment to ethnically homogenous groups, e.g. where participants reported themselves as white with all grandparents were born in the UK. To account for more subtle influences of population stratification, principal component analysis is frequently used to construct covariates capturing correlations of SNPs across the genome. These have been shown to clearly capture subtle ancestry differences with high accuracy [Novembre and others 2008]. The second source of false positives is more intrinsic to the nature of GWAS, with the testing of hundreds of thousands of markers leading to a very large number of SNPs expected to be significant at $p<0.05$ by chance (since this reflects a 1 in 20 chance of the result occurring at random). To account for this a genome-wide significance level is often implemented at p=5E-8, estimated to account for the number of tests across the genome [Dudbridge and Gusnanto 2008]. This high threshold for significance means large sample

sizes are required to have adequate power to detect true effects, and a lower threshold is often used to identify SNPs at a suggestive significance to be taken forward for replication.

1.3.2.2 Findings

The initial GWAS in psychiatric disorders showed little to no success in identifying causal risk variants, unlike some somatic disorders [Wellcome Trust Case Control Consortium 2007].There were several exceptions to this, mostly in those psychiatric disorders where the biology was better understood. As mentioned the first GWASs into Alzheimer's disorder quickly identified novel findings [Harold and others 2009; Hollingworth and others 2011; Lambert and others 2009; Naj and others 2011] despite relatively small sample sizes, including regions originally identified in candidate studies. GWAS of nicotine addiction also identified variants in nicotine receptor genes [Tobacco and Genetics Consortium 2010; Liu and others 2010; Thorgeirsson and others 2010] and for alcohol usage and dependence in alcohol dehydrogenase genes previously associated with ability to break down alcohol [e.g. Bierut and others 2012; Gelernter and others 2014].

While the initial disappointing results for most disorders was taken by some as evidence against the common disease common variants hypotheis, efforts were made by researchers to pool samples together into meta-analyses (or mega-analyses, with data analysed together) in order to increase power to detect SNPs of small effect [e.g. Purcell and others 2009]. This approach came to fruition most clearly with the Psychiatric GWAS Consortium (later the Psychiatric Genomics Consortium or PGC), which in 2011 published two GWAS on schizophrenia and bipolar disorder [Ripke and others 2011; Sklar and others 2011]. The schizophrenia analysis combined 9,394 cases and 12,462 controls from 17 studies and identified 7 risk loci for schizophrenia, including the major histocompatability complex (MHC) region that had previously been suggested in earlier candidate gene studies [McGuffin 1979]. The analysis of bipolar disorder identified two genome-wide significant loci in a sample of 7,481 cases and 9,250 controls from 12 studies. Both studies were able to show that among the top associations, there was a highly significant excess of SNPs where

the direction of effect was concordant within the discovery and independent replication datasets suggesting that with greater sample sizes some of these SNPs might reach genome-wide significance. A future meta-analysis of the PGC schizophrenia studies with several thousand additional cases and controls from Swedish samples would prove this true [Ripke and others 2013a], identifying 13 more loci at genome-wide significance. A similar trend was seen in Alzheimer's disease, with the pooling of samples into the GERAD consortium of 17,008 cases and 37,154 controls to identify 20 associated loci [Lambert and others 2013], and as previously mentioned GWAS to some extent verified previously known variants from linkage studies. These results showed that once large enough sample sizes had been achieved, GWAS could identify novel associations and insights into the aetiology of psychiatric disorders.

Not all large GWAS collaborations yielded such successes though. The PGC analysis of MDD failed to identify any genome-wide significant associations, and the concordance between discovery and replication samples for the top SNPs was much weaker, despite a similar sample size of 9,240 cases and 9,519 controls [Ripke and others 2013b]. Similar negative findings were found for ADHD and autism [Neale and others 2010; Smoller and others 2013], though here the sample sizes were smaller. However the success in schizophrenia motivated the collection of genotyped samples for new disorders, both within and without the PGC, including anorexia nervosa [Boraska and others 2014; Wang and others 2011], cannabis use [Verweij and others 2013], Tourette's [Scharf and others 2013], OCD [Stewart and others 2013] , anti-social behaviour disorder [Tielbeek and others 2012], anxiety [Trzaskowski and others 2013], insomnia [Byrne and others 2013] and borderline personality [Lubke and others 2013]. Though most of these GWAS led to negative findings, a few successes stand out, namely post-traumatic stress disorder [Xie and others 2013] and in a very large meta-analysis of educational attainment [Rietveld and others 2013], as a proxy for cognitive functioning. Despite the current abundance of negative findings, increasingly large samples seem likely to lead to a greater and greater number of associations

1.3.3 Polygenic methods

For the most part even where genome-wide significant associations have been identified in GWAS, they still explain only a very small proportion of the variance in risk for psychiatric disorders. Hence several methods have been designed to make use of GWAS data to estimate the amount of variance explained when the effects of variants across the genome are combined. One of these was alluded to briefly in the above section, whereby concordance between the direction of effect in the discovery sample's top SNP associations is compared to the direction of effect in the replication sample (known as a sign test). However most of the methods discussed here focus not on the top SNPs from a GWAS but on all available SNPs. This is due to the stringent threshold for genome-wide significance and the presumption that many truely causal SNPs of small effect might not meet it, yet their cumulative effect could explain a large amount of the variance in risk.

1.3.3.1 Polygenic risk profile scoring

The first method to be discussed is polygenic risk profile scoring (PRS). This is similar to a sign test in that it requires a discovery sample and a replication (or target) sample. In the discovery sample a GWAS is performed, defining the risk allele and effect size for each SNP. In the target sample the number of risk alleles an individual carries is summed into a score, with each SNP weighted by its effect size. This is then used as a predictor of risk to determine how much of the variance between cases and controls is explained by the PRS. The PRS is often defined using various p-value thresholds for the inclusion of SNPs from the discovery GWAS, e.g. four scores using SNPs with $p<0.001$, $p<0.05$, $p<0.2$, and $p<0.5$. As more SNPs are included, the greater the likelihood that all SNPs of true effect will be captured. However including more SNPs also includes more with no true effect and so adds noise, causing a levelling off of the amount of variance explained. The first application of PRS was in the International Schizophrenia Consortium, which showed ~2% of the variance could be explained across samples [Purcell and others 2009]. They were also able to show that PRS for schizophrenia could also predict bipolar disorder vs. controls with a slightly reduced effect size. Future studies within the PGC would show that about 7% of the variance for

schizophrenia was predicted by PRS [Ripke and others 2011], 3% for bipolar disorder [Sklar and others 2011], and 0.6% for MDD [Ripke and others 2013b]. Cross disorder analysis also showed that schizophrenia, bipolar disorder and MDD PRS all predicted across each other, with some more weakly predicting into autism and ADHD [Smoller and others 2013]. More complex Bayesian approaches to PRS have also estimated the number of independent risk variants for schizophrenia to be between 6,300–10,200, accounting for 32% of the variance in risk [Ripke and others 2013a].

1.3.3.2 Genomic-relatedness-matrix residual maximum likelihood analysis

An alternate polygenic method, genomic-relatedness-matrix residual maximum likelihood analysis (GREML), was designed not for prediction of risk but to estimate what proportion of variance is captured by SNPs, thus giving a heritability estimate based on known genetic variants rather than from twin models [e.g. GCTA, developed by Yang and others 2011]. This method works by estimating the relatedness of all pairs of individuals in the dataset, and testing how much of the similarity in phenotype within pairs is accounted for by genetic similarity (relatedness). Relatedness here is defined as the proportion of the genome for which they have the same genotypes. It is important to note that the results of this analysis do not reflect the true heritability of a trait, as it only includes that heritability captured by the included SNPs while usually ignoring the role of rare or structural variation. It does however act as a benchmark for the maximum amount of risk detectable in an infinitely powered GWAS. An analysis within the PGC showed schizophrenia, bipolar disorder, MDD, ADHD, and autism all had relatively similar SNP-heritabilities at around ~20-25%, with the highest in ADHD (28%) and the lowest in autism (17%) [Lee and others 2013b]. GREML analysis can also be applied using bivariate models to estimate the co-heritability of two disorders captured by SNPs, with the key benefit of not requiring individuals to have data on both phenotypes [Lee and others 2012]. MDD, schizophrenia and bipolar disorder showed highly significant co-heritabilities at around 10-15%, as well as slightly lower co-heritabilities for MDD-ADHD and schizophrenia-autism [Lee and others 2013b]. Secondary

analyses suggested that while a large proportion of the heritability consistently came from genes annotated as related to brain function, it also unexpectedly showed that within disorders different studies did not overlap in their genetics entirely e.g. the heritability of MDD in study 1 might only show a co-heritability of 60% with the heritability of MDD in study 2, presumably due to heterogeneity across studies. Other studies have shown significant heritabilities for Alzheimer's (24%) [Lee and others 2013a] and Tourette's and OCD (58% and 37%, with significant genetic correlation between them) [Davis and others 2013]. Other studies have shown that SNP-heritability substantially overlaps across populations [de Candia and others 2013; Yang and others 2013]. These findings broadly are in agreement with those from PRS studies, suggesting that common risk variants contribute substantial risk to psychiatric disorders despite the individual risk variants not yet being identified. Presumably this is due to a large number of variants of small effect, with the results suggesting that these variants contribute non-specific risk to psychiatric disorders in general rather than uniquely to one disorder.

1.3.3.3 Runs of homozygosity

The last polygenic method discussed here is the analysis of runs of homozygous markers across the genome. An excess of homozygosity across the genome reflects inbreeding, whether through more recent events, such as consangous marriages, or among more distant ancestors. It has been shown that by looking for sections of consecutive homozygous markers more accurately captures inbreeding over using simply percentage of genome is homozygous [Keller and others 2011]. Homozygosity not only reflects inbreeding but also exposes recessive variants, those that have a non-additive effect with number of alleles an individual carries. A fully recessive variant is only expressed when an individual carries a copy on both chromosomes, though variants with partially recessive effects exist. As this can hide deleterious alleles from being expressed, natural selection acts to increase the frequency of dominant beneficial alleles and selects out more quickly dominant deleterious alleles. As a result, one expects deleterious traits to show an excess of recessive alleles.

Hence if these runs of homozygosity (ROH) are associated with a trait or disorder, it is evidence that the disorder has experienced historical selection against it. This informs both on its negative effect on ancestral human reproductive fitness, and that a recessive model of risk may be more appropriate when attempting to identify risk variants. Such an association has been suggested by several studies of psychiatric disorders, most prominently in schizophrenia [Keller and others 2012], intellectual disability and autism [Gamsiz and others 2013]. However, due to the very small effect sizes expected future studies require very large sample sizes in order to identify such an association [Keller and others 2011].

1.3.5 Whole Genome Sequencing

In the last few years GWAS have been followed up by whole-genome sequencing (WGS). This approach has become more popular as the cost has decreased and next generation technologies have been developed. While GWAS relies on using common genetic markers to tag local variation, WGS allows for the identification of the causal variant directly. It also allows for the analysis of rare variants, in particular *de novo* mutations that could never be found on a genotyping chip of known variants. The downside of this approach is that the sheer number of variants to test increases dramatically. This has several draw backs: by increasing the computational burden of analyses; by inflating the burden of multiple testing and so the threshold for significance; by including not just common variants but also rare variants where the power to detect significance is greatly reduced; and by not restricting to known variants and so increasing the risk of a incorporating sequencing errors. As such many methods for analysing WGS data have focused on analysing burden of rare variants across segments of the genome, or have returned to the format of linkage studies by leveraging pedigree information to investigate co-segregation of disease and variant. These methods often have led to a focus on variants within genes, where their impact on the protein can be estimated and weighted for. As such some studies have focused on exome sequencing, targeting only the portions of the genome that transcribe to mRNA. Within psychiatric disorders, Alzheimer's

again has led the way with the discovery of several variants in exome and genome sequencing studies [Guerreiro and others 2012; Pottier and others 2012]. And within other psychiatric disorders, the results are again less impressive to date. The most recent large scale analysis of 2,536 schizophrenia cases and 2,543 controls' exome sequences suggested cases showed a polygenic burden of disruptive mutations, particularly concentrated in previously associated gene pathway [Purcell and others 2014]. Several studies of exome sequencing in autism came out in 2012, implicating rare and *de novo* mutations though unable to identify individual loci definitively [Iossifov and others 2012; Neale and others 2012; O'Roak and others 2012; Sanders and others 2012]. These studies, as well as for schizophrenia [Fromer and others 2014], found an excess of *de novo* mutations amongst cases, particularly those with later paternal ages, providing evidence that these novel mutations may play a crucial role in psychiatric disorders.

1.3.6 Structural variation

As well as point mutations in the genome, where one base pair is replaced by another, there is the possibility of structural variation where base pairs or segments of the genome are deleted or duplicated. While many forms of structural changes exist, current genetic research has focused on larger copy number variants (CNV) that can be *de novo* or remain in the population as CNV polymorphisms. Schizophrenia has been seen as highly prevalent within the syndromic CNVs, such as 22q11 cardiofacial syndrome. As well as these syndromic effects, several large and rare CNVs of large effect are known to convey risk to schizophrenia, though also numerous other disorders such as autism, intellectual disability and epilepsy [reviewed inSullivan and others 2012]. Intellectual disability and autism both have a high burden of large structural variants or in a large number of cases [Folstein and Rosen-Sheidley 2001; Sanders and others 2011; Vorstman and others 2006] , sometimes established as the cause such as in the case of fragile X syndrome. As well as the identification of large CNVs, GWAS data can be used to identify smaller CNVs. This is done by looking at the signal intensity for the alleles, and looking for stretches of homozygosity or a 2:1

ratio of heterogeneity. These methods have been used to look at burden of CNVs across the genome for various psychiatric disorders. Such an association established in several studies of schizophrenia [Szatkiewicz and others 2014; Walsh and others 2008; Xu and others 2008], ADHD [Williams and others 2012; Williams and others 2010; Yang and others 2013] and autism [Pinto and others 2010], particularly among de novo CNVs and those in genic regions [Kirov and others 2012; Szatkiewicz and others 2014]. The results for bipolar disorder are more mixed [Grozeva and others 2010; Malhotra and others 2011; McQuillin and others 2011; Priebe and others 2012; Zhang and others 2009], while only one association for MDD with CNV burden has been reported [Rucker and others 2013]. Alzheimer's again shows concordance across different methods with associations with variants across the APP region [McNaughton and others 2012; Rovelet-Lecrux and others 2006].

1.4 Summary and outline of thesis

1.4.1 State of psychiatric genetics findings

These molecular and epidemiological findings have started to give insight into the genetic architecture of psychiatric genetics. While individual risk variants are still relative few for most disorders, a few patterns have started to emerge. Structural, rare and common variants for schizophrenia, bipolar disorder and autism appear to cluster in genes with brain function. Polygenic approaches have shown that common variants account for around ~15-30% of the variance in psychiatric disorders, approximately a third or half of the heritability. This suggests that as GWAS studies increase in size they will have sufficient sample size to identify many more common variants of small effect. Studies of cumulative burden of rare and *de novo* structural and point mutations across the genome also suggest that larger samples will yield associations with individual variants. Epidemiological evidence for paternal age effects suggests these approaches may have the greatest success for schizophrenia and autism. However progress may depend on the speed at which whole genome sequencing becomes viable, as well as the development of

methodologies for analysing very rare variants. Hopefully the disorder with the greatest success, Alzheimer's disease, will act as a template whereby rare, common and structural variants converge onto specific regions, presenting biological insight and drug targets.

This success will no doubt come through the ability to generate larger and larger sample sizes for a variety of disorders, something that is already occurring through cheaper genotyping/sequencing technologies and the acceptance of the need for international collaborations. In the meantime, several approaches can make use of existing data to gain a better picture of the underlying genetics of psychiatric disorders. My thesis will make use of six papers across three areas of research to gain novel insight into the genetic architecture of psychiatric disorders. The following sections will outline the contents and objectives of each.

1.4.2 Impact on reproductive fitness

In the first section of my thesis I will examine two ways in which psychiatric disorders can impact reproductive fitness and behaviour. The first paper, "Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings" [Power and others 2013b], focuses on a systematic analysis of the impact of psychiatric disorders on fecundity, as measured by number of children in both those diagnosed with a disorder and unaffected relatives. As discussed above this has huge impact on how causal variants are maintained in the population, and so the distribution of frequencies and effect sizes. The second paper "A recessive genetic model and runs of homozygosity in major depressive disorder" focuses on major depression following the results of the previous paper, looking for molecular support for a lack of negative selection on depression in the Psychiatric Genomics Consortium's sample of over 9000 each of MDD cases and controls with genome-wide association data [Power and others 2014a]. Together they paint a picture of where we might expect to identify causal variants in molecular genetic studies, and how these expectations differ across psychiatric disorders.

1.4.3 Gene-environment interactions

Following the results of the previous section showing potential evidence for beneficial effects of genetic risk variants for MDD, the next section will focus on gene-environment interactions and whether they could provide an answer to why these variants appear to sometimes increase an individual's fitness. It is clear that some individuals succumb to environmental risk factors while others do not. Whether risk variants for a disorder predispose individuals only in certain environments seems plausible, though the evidence for such gene-environment interactions is often highly contested.

In this section I first test the 'Mismatch hypothesis' which suggests environmental risks may interact with each other differently based on an individual's genetic background of 'plasticity' variants in a paper titled "The interaction between child maltreatment, adult stressful life events and the 5-*HTTLPR* in major depression" [Power and others 2013c]. In a second paper, I look not at testing for gene-environment interactions but to stratify cases and controls into equivalent levels of environmental stress exposure, in a paper "Genome-wide association analysis accounting for environmental factors through propensity-score matching: Application to stressful live events in major depressive disorder" [Power and others 2013a]. Both approaches aim to move beyond the conventional gene-environment interaction analysis, which has to date produced limited returns.

1.4.4 Untangling genetic from environmental risk

Following from the findings of the previous section, I explore how the relationship between environmental and genetic risk factors is complicated by the possibility not just of interactions but also correlations. There are several ways in which this can occur, with individuals selecting their own preferred environments and having their environments react to them. This can contribute to gene-environment correlation, particularly with respect to behavioural traits, and creates a problem for disentangling the causation of associations in the epidemiology of psychiatric trait. In two papers, "Estimating the heritability of reporting stressful life events captured by common

genetic variants" [Power and others 2013d] and "Genetic predisposition to schizophrenia associated with increased use of cannabis" [Power and others 2014b], I test for an underlying heritability for the known 'environmental' psychiatric risk factors of stressful life events and cannabis use, and look at how this heritability may overlap with the heritability of psychiatric disorders themselves.

1.4.5 Conclusion

Psychiatric genetics is on the verge of major advances, largely due to the massive increase in available data. As yet there are still many unanswered questions about the underlying aetiology of these disorders, such as their impact on reproductive fitness, the extent to which we can dissect heterogeneity within disorders through genetics, and the complex relationships between genes and environment. Overall this thesis will combine epidemiological, genetic and environmental data to attempt to answer some of these questions.

ORIGINAL ARTICLE

Fecundity of Patients With Schizophrenia, Autism, Bipolar Disorder, Depression, Anorexia Nervosa, or Substance Abuse vs Their Unaffected Siblings

Robert A. Power, BSc; Simon Kyaga, MD; Rudolf Uher, MD, PhD, MRCPsych; James H. MacCabe, PhD, MRCPsych; Niklas Långström, MD, PhD; Mikael Landen, MD, PhD; Peter McGuffin, FRCP, FRCPsych, PhD; Cathryn M. Lewis, PhD; Paul Lichtenstein, PhD; Anna C. Svensson, PhD

Context: It is unknown how genetic variants conferring liability to psychiatric disorders survive in the population despite strong negative selection. However, this is key to understanding their etiology and designing studies to identify risk variants.

Objectives: To examine the reproductive fitness of patients with schizophrenia and other psychiatric disorders vs their unaffected siblings and to evaluate the level of selection on causal genetic variants.

Design: We measured the fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse and their unaffected siblings compared with the general population.

Setting: Population databases in Sweden, including the Multi-Generation Register and the Swedish Hospital Discharge Register.

Participants: In total, 2.3 million individuals among the 1950 to 1970 birth cohort in Sweden.

Main Outcome Measures: Fertility ratio (FR), reflecting the mean number of children compared with that of the general population, accounting for age, sex, family size, and affected status.

Results: Except for women with depression, affected patients had significantly fewer children (FR range for those with psychiatric disorder, 0.23-0.93; $P<10^{-10}$). This reduction was consistently greater among men than women, suggesting that male fitness was particularly sensitive. Although sisters of patients with schizophrenia and bipolar disorder had increased fecundity (FR range, 1.02-1.03; $P<.01$), this was too small on its own to counterbalance the reduced fitness of affected patients. Brothers of patients with schizophrenia and autism showed reduced fecundity (FR range, 0.94-0.97; $P<.001$). Siblings of patients with depression and substance abuse had significantly increased fecundity (FR range, 1.01-1.05; $P<10^{-10}$). In the case of depression, this more than compensated for the lower fecundity of affected individuals.

Conclusions: Our results suggest that strong selection exists against schizophrenia, autism, and anorexia nervosa and that these variants may be maintained by new mutations or an as-yet unknown mechanism. Bipolar disorder did not seem to be under strong negative selection. Vulnerability to depression, and perhaps substance abuse, may be preserved by balancing selection, suggesting the involvement of common genetic variants in ways that depend on other genes and on environment.

JAMA Psychiatry. 2013;70(1):22-30.
Published online November 12, 2012.
doi:10.1001/jamapsychiatry.2013.268

Author Affiliations are listed at the end of this article.

PSYCHIATRIC DISORDERS HAVE long puzzled researchers by defying the expectations of natural selection.[1] From an evolutionary viewpoint, selection should remove genetic variants that reduce an individual's ability to reproduce ("fitness") because they will produce fewer offspring to inherit those variants. However, psychiatric disorders do not fit this model. They combine substantial heritability with moderate to high prevalence, early age at onset, and reduction in fitness compared with the general population. Several hypotheses have been put forward on how psychiatric disorders survive in the population,[2-5] but the mechanisms that maintain the genetic variants conferring susceptibility to these disorders remain unclear.

Insights into the form of selection pressure on psychiatric disorders may help direct research toward identifying specific causal pathways. The large disparity between heritability estimates from family studies and the amount of variance in psychiatric disorders explained by identified risk alleles has become known as the "missing heritability."[6] Current genome-wide association studies are based on the

common disorder–common variants hypothesis, which presumes that many low-risk high-frequency alleles directly lead to highly prevalent complex disorders. Evidence suggesting an alternative reason why risk alleles remain in the population could explain why more causal variants have not been discovered through linkage or association studies.

Several alternatives to the common disorder–common variants hypothesis exist. Autism, bipolar disorder, and schizophrenia have been associated with increased paternal age.[7-10] Older paternal age carries a risk of an increased number of de novo mutations during spermatogenesis,[11] which in turn may lead to deleterious phenotypes in the next generation.[12] Alternatively, deleterious genes that exist in the population may only recently have come under purifying selection, as changes in selection pressures have made them detrimental in the present environment. A competing hypothesis is that causal genetic variants may not be entirely deleterious but may also confer benefits (eg, creativity[13]) and that balancing selection maintains these variants in the population at optimal frequency. Possible mechanisms for balancing selection include heterozygote advantage, pleiotropic antagonism, and gene-environment interactions. A specific form of balancing selection is sexual antagonism, where a trait may be beneficial to one sex but harmful to the other.

To our knowledge, this analysis is the first to evaluate the fecundity of affected individuals and their siblings for multiple psychiatric disorders, including schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, and substance abuse. Previous research has focused solely on psychosis, and a recent meta-analysis[14] found no overall increase in the fecundity among siblings of those with schizophrenia. Similar evidence is missing for most other psychiatric disorders. The aim of this study was to examine sibling fecundity to evaluate evidence for the aforementioned theories for the continued prevalence of psychiatric disorders (**Table 1**). Under balancing selection, we expected that unaffected siblings of those with psychiatric disorders would enjoy increased reproductive fitness because they would benefit from the positive effects of the same genetic variants that contributed to psychiatric disorders in their kin. Sex-specific effects on sibling fecundity may provide evidence for sexual antagonism, with the same genetic variants benefiting one sex at the cost of the other. For de novo or recent highly penetrant mutations, no reduction in the fecundity would be expected in unaffected siblings (as long as affected siblings are diagnosed). If these disorders were affected by multiple variants that were ancestrally neutral but deleterious in today's environment, we expected to see decreased fecundity in siblings due to an increased probability of sharing those variants and the unsuited environment.

Table 1. Hypotheses of Maintenance of Genes for Psychiatric Disorders and Predicted Results for Fecundity in Affected Individuals and Their Unaffected Siblings

	Predicted Result for Fecundity	
Hypothesis	Affected Individuals	Unaffected Siblings
Balancing selection	Decreased	Increased in proportion to the decrease in affected individuals
Sex-dependent selection	Decreased to a greater extent in one sex	Decreased in the same sex as that of affected individuals but increased in the opposite sex
De novo mutation	Greatly decreased	No change
Ancestral neutrality	Decreased	Decreased to a lesser extent owing to shared genes and environment

METHODS

POPULATION DATABASES

The data set was drawn from the Multi-Generation Register, which includes children born in Sweden since 1932 and their biological parents.[15] All individuals alive in 1960 and all births from this point onward were recorded in the register. Paternity is assumed to be the husband of the mother at the time of birth or "by acknowledgment" for unwed mothers. This data set was linked to the Swedish Hospital Discharge Register, which covers virtually all psychiatric hospitalizations since 1973 in Sweden[16] and has been previously validated.[17,18] It also includes partial coverage of outpatient diagnoses from 2001 onward. The use of this database has been approved by the ethics committee at the Karolinska Institutet, Stockholm, Sweden. Diagnoses were established according to the *International Classification of Diseases* (*ICD*), *Eighth Revision* (1973-1987), *Ninth Revision* (1987-1996), and *Tenth Revision* (1996 onward). These registers were linked using an individual's unique national registration number. The analyses were restricted to 2 356 598 individuals born in Sweden between 1950 and 1970, for whom most of their adult life was covered in the Swedish Hospital Discharge Register and who would have completed most of their reproductive life span (age 40 years for the youngest individuals). Only individuals for whom both parents were known were included in the birth cohort, although this limitation was not applicable when calculating the number of offspring.

DISORDERS CLASSIFICATION

Six disorders were examined, including schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, and substance abuse (including alcohol use disorder). The 6 disorders were chosen to differ in terms of prevalence, severity, heritability, and their distribution between men and women. In all cases, previous studies[19-24] using the Multi-Generation Register and the Swedish Hospital Discharge Register were used as a basis for the selection of *ICD* codes. Following the validation of bipolar disorder in the Swedish Hospital Discharge Register by Sellgren et al,[24] 2 or more diagnoses over a lifetime were required o be included. For all other disorders, only a single lifetime diagnosis in the Swedish Hospital Discharge Register was required. Svensson et al[23] had previously examined schizophrenia in the Multi-Generation Register and the Swedish Hospital Discharge Register in an earlier (nonoverlapping) birth cohort. Schizophrenia was defined by 1 or more recorded diagnoses of code 295 in *ICD-8* and *ICD-9* and by codes F20, F23.1, F23.2, and F25 in *ICD-10*. Autism was defined by code 299.0 in *ICD-9* and by codes F84.0, F84.1, F84.5, and F84.9 in *ICD-10* and was not defined in *ICD-8*. Bipolar disorder was defined by code 296 (excluding code 296.2) in *ICD-8*, by code 296 (excluding code 296.1) in *ICD-9*, and by codes F30 and F31 in

Table 2. Epidemiological Details for 6 Psychiatric Disorders Among the Cohort of 2 356 598 Individuals Born in Sweden Between 1950 and 1970

Psychiatric Disorder	No. of Affected Individuals	Prevalence, %	Female-Male Ratio	No. of Children Among the Cohort, Mean No. of Affected	No. of Unaffected Siblings
Schizophrenia	18 890	0.80	1:1.5	0.61	28 644
Autism	2947	0.12	1:2	0.56	4471
Bipolar disorder	14 439	0.61	1.5:1	1.48	22 986
Depression	81 295	3.44	1.4:1	1.78	119 645
Anorexia nervosa	3275	0.14	10.4:1	1.46	5172
Substance abuse	55 933	2.37	1:2.3	1.49	81 592

ICD-10. One of these was allowed to be code 296.2 in *ICD-8* or code 296.1 in *ICD-9*. Depression was defined by codes 296.0, 296.2, and 298.0 in *ICD-8*; by codes 296.2, 296.3, and 298.0 in *ICD-9*; and by codes F32 and F33 in *ICD-10*. Anorexia nervosa was defined by code 306.5 in *ICD-8*, by code 307.1 in *ICD-9*, and by code F50.0 in *ICD-10*. Last, substance abuse was defined by codes 303 and 304 in *ICD-8*; by codes 303, 304, 305.1, and 305.9 in *ICD-9*; and by codes F10 through F19 (excluding subsection 0.5, diagnosing substance abuse with psychosis) in *ICD-10*. Diagnoses were made on discharge by the treating physician. No hierarchical diagnostic practice was used; hence, individuals with comorbidity could appear in more than 1 category. A large amount of comorbidity existed within the sample, with affected individuals being diagnosed as having at least 1 other disorder in 45.7% of individuals with schizophrenia, 48.9% with autism, 71.7% with bipolar disorder, 26.4% with depression, 26.8% with anorexia nervosa, and 30.0% with substance abuse. High levels of comorbidity between psychiatric disorders is a common feature of cohort studies and has been described at length elsewhere.[25] For the objectives of this study, we first analyzed each disorder separately without accounting for comorbidities. A secondary analysis was then performed that corrected for comorbidities by analyzing all disorders simultaneously. Family identification codes were used to identify unaffected full siblings, while affected siblings were considered affected individuals, not siblings. The results for our identification of affected individuals and their siblings are summarized in **Table 2**.

MEASURE OF FECUNDITY

Data were analyzed using generalized estimating equations,[26,27] accounting for similarity in the number of children within families. To measure the reproductive fitness of each group, a fertility ratio (FR) was calculated based on the number of children individuals in that group had compared with the general population, correcting for the year of birth. For example, if the disease group had an FR of 0.5, it meant they had on average half as many children as the general population, while an FR of 2 meant they had twice as many. To permit testing for sex-specific effects and to avoid confounding by age differences at parenthood and the mean number of children, we compared affected men with the general population of men, and the same for affected women. Similarly, siblings of affected individuals were compared with only those in the general population of the same sex, with the additional requirement that they had at least 1 sibling, to account for any bias resulting from being a sibling. Socioeconomic effects have been shown to influence the risk for psychiatric disorders[28] and could potentially confound our analysis. To account for socioeconomic status, we corrected for both paternal and maternal education levels derived from the 1970 census data.[29] Parental education level rather than the self-education level was used to avoid reverse causation. Data on at least 1 parent were available for more than 95% of individuals.

To interpret the FRs of affected individuals and their siblings together, we followed the method by Haukka et al[30] and compared the prevalence of affected individuals and siblings plus that of their combined children. Therefore, we combined the estimated number of children from affected individuals and their siblings, and we then divided that sum by the estimated total number of children for the entire 1950 to 1970 birth cohort. The total estimated number of children was derived from the FR of each group (affected, sibling, and remaining population), multiplied by the mean number of children, and weighted for by each group's frequency in the birth cohort. Using the FR rather than the actual number of children of each individual, we corrected for the year of birth and the unequal distribution of affected individuals born each year. For this analysis, we used the FRs that were not corrected for comorbidities. All the analyses were performed using commercially available software (STATA, release 12; StataCorp LP).[31]

RESULTS

The mean (SD) number of children for the birth cohort was 1.76 (1.27). Paternal and maternal education levels were found to be significantly associated with fecundity. Because only minor differences were observed between the adjusted and nonadjusted estimates, only the adjusted estimates are presented herein. The mean (SD) number of siblings for the birth cohort was 1.68 (1.38).

SCHIZOPHRENIA

Individuals with schizophrenia had fewer children compared with the general population, with FRs of 0.23 (95% CI, 0.23-0.24; $P < 10^{-10}$) for men and 0.47 (95% CI, 0.46-0.48; $P < 10^{-10}$) for women (**Figure 1**). Sisters of affected individuals had a significantly increased number of children (FR, 1.02; 95% CI, 1.01-1.03; $P = .01$), while brothers of affected individuals showed significantly decreased fecundity (FR, 0.97; 95% CI, 0.96-0.99; $P < .001$) (**Figure 2**). When comorbidities were included in the analysis, the increased fecundity in sisters disappeared.

AUTISM

Individuals with autism had significantly fewer children, in men (FR, 0.25; 95% CI, 0.23-27; $P < 10^{-10}$) and women (FR, 0.48; 95% CI, 0.44-0.51; $P < 10^{-10}$). Brothers of affected individuals also had fewer children (FR, 0.94; 95%

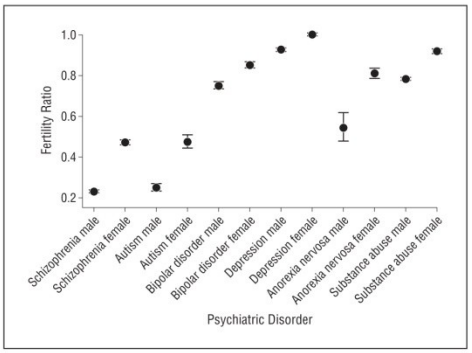


Figure 1. Fertility ratios for individuals with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, and substance abuse. A fertility ratio of 1 (highlighted) represents that of the general population.

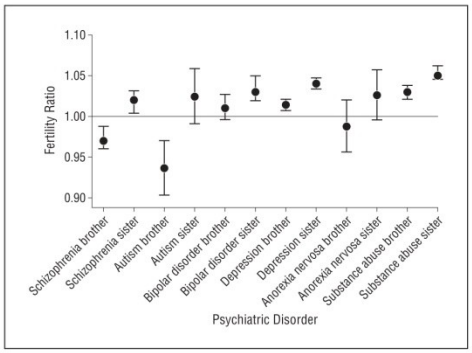


Figure 2. Fertility ratios for unaffected siblings of individuals with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, and substance abuse. A fertility ratio of 1 (highlighted) represents that of the general population.

CI, 0.90-0.97; $P < .001$), while sisters of affected individuals showed no significant difference from the general population. These results did not differ significantly when comorbidities were included in the analysis.

BIPOLAR DISORDER

Men and women with bipolar disorder had fewer children than the general population (male FR, 0.75; 95% CI, 0.73-0.77; $P < 10^{-10}$; female FR, 0.85; 95% CI, 0.84-0.87; $P < 10^{-10}$). Brothers of affected individuals showed no significant difference from the general population, while sisters of affected individuals had an increased number of children (FR, 1.03; 95% CI, 1.02-1.05; $P < 10^{-5}$). When correcting for comorbidity, the increased fecundity in sisters disappeared, and the reduced fecundity in affected individuals increased to just below that of the general population (male FR, 0.94; 95% CI, 0.92-0.96; $P < .001$; female FR, 0.95; 95% CI, 0.93-0.97; $P < .001$).

DEPRESSION

Men with depression had fewer children (FR, 0.93; 95% CI, 0.92-94; $P < 10^{-10}$), but women with depression showed no significant difference from the general population. Siblings of affected individuals had more children compared with the general population (brothers' FR, 1.01; 95% CI, 1.01-1.02; $P < .0001$; sisters' FR, 1.04; 95% CI, 1.03-1.05; $P < 10^{-10}$). This increased fecundity in siblings remained when comorbidities were accounted for, although the reduced number of children among affected men disappeared and the fecundity among women with depression increased (FR, 1.03; 95% CI, 1.03-1.04; $P < .001$).

ANOREXIA NERVOSA

Individuals with anorexia nervosa showed a reduced number of children in men (FR, 0.54; 95% CI, 0.48-62; $P < 10^{-10}$) and in women (FR, 0.81; 95% CI, 0.79-84; $P < 10^{-10}$). Neither the fecundity of brothers nor sisters differed from that of the general population. These results did not change after correction for comorbidities.

SUBSTANCE ABUSE

Men having a diagnosis of substance abuse had significantly fewer children than the general population (FR, 0.78; 95% CI, 0.78-79; $P < 10^{-10}$), as did women having a diagnosis of substance abuse (FR, 0.92; 95% CI, 0.91-0.93; $P < 10^{-10}$). Siblings of individuals with substance abuse had more children than the general population, with FRs of 1.03 (95% CI, 1.02-1.04; $P < .0001$) for brothers and 1.05 (95% CI, 1.05-1.06; $P < 10^{-10}$) for sisters. These values did not differ significantly when comorbidities were included.

COMMENT

To our knowledge, this is the first time the fitness of relatives has been examined for individuals with these psychiatric disorders (other than schizophrenia and bipolar disorder). Across disorders, affected men had a consistently greater reduction in fecundity than affected women. A similar sex difference was observed among siblings: sisters of individuals with psychiatric disorders had more children than their brothers. The degree of fecundity reduction in affected individuals and the associated increase in the fecundity among siblings differed by disorder, which suggests that different types of psychiatric disorders are under different selection pressures.

STUDY LIMITATIONS

Our conclusions rely on the assumption that the fecundity measured truly reflects these individuals' reproductive fitness. Ideally, the number of grandchildren rather than children would be used to measure an individual's long-term fitness, and this should become possible in the future with the continuation of the Multi-Generation Register in Sweden. The fecundity in affected individuals may be decreased by effects of medication or hospitalization, which may have distorted our findings. Voluntary and compulsory sterilizations were historically performed in Sweden until the 1970s, including targeting those with

Table 3. Fitness Calculations of the Observed Proportions of Affected Individuals in the 1950 to 1970 Birth Cohort and the Estimated Proportions of Their Affected Children in the Next Generation

Psychiatric Disorder	Proportion of Affected Individuals		
	Observed in the 1950-1970 Birth Cohort	Estimated in the Next Generation	Change, %
Schizophrenia	0.022	0.016	−25
Autism	0.003	0.002	−28
Bipolar disorder	0.012	0.011	−6
Depression	0.092	0.093	1
Anorexia nervosa	0.003	0.002	−10
Substance abuse	0.066	0.064	−3

psychiatric disorders. It should be noted that most of those sterilized were women, while our results consistently show a greater reduction in male fecundity.[32,33] Last, children of those with a psychiatric disorder may go unrecorded as a result of stigma or chaotic lifestyles (eg, emigration), leading to an artificial lowering of their fecundity. However, these concerns are not applicable to unaffected siblings. Therefore, we believe that the number of children born to siblings of individuals with psychiatric disorders accurately reflects their fitness.

Is reproductive success today representative of reproductive success in the past? Recent reductions in child mortality and in the mean number of children per adult, as well as the increased use of contraceptives, raise questions about how well evolutionary fitness can be measured in the modern world. This is important because a trait's prevalence is dependent on the selection it has encountered in previous generations, not the current one. However, several reasons indicate why we would expect fitness today to reflect fitness in previous generations to some extent. First, changes in fecundity may reflect a biological or physiological difference in fertility that affects fitness regardless of culture or setting (eg, reduced sperm count). Second, it has been argued that the inability to attract a mate is responsible for low fitness in those with psychiatric disorders (reviewed by Keller and Miller[2]), supported by evidence showing that low marriage rates mediate the effect of psychiatric disorders on fitness.[34] If this is the case, then modern improvements in contraceptives and child mortality occur after the selection against individuals with psychiatric disorders has already occurred and so have less influence. Third, the extent of the impairment experienced by those with psychiatric disorders in traditional communities is high, suggesting that the effects are not culture specific.[35,36] Although our measure was suboptimal, we maintain that the large differences in fecundity suggest as close a measure of fitness as any other available in human data sets.

Finally, the generalization of a single etiology to a disorder may be unfounded. As described in the "Methods" section, considerable comorbidity exists between disorders. Longitudinal, family, and molecular studies[25,37,38] have demonstrated that psychiatric disorders show considerable overlap and share genetic risk variants. We attempted to tackle this by analyzing performing analyses with vs without accounting for comorbidities, and this had little effect (except for bipolar disorder). More problematic is the evidence for considerable genetic and phenotypic heterogeneity within psychiatric disorders.[39] Heterogeneity within disorders means that caution must be applied to generalizations of their genetic etiology.

SEX-SPECIFIC EFFECTS

Across disorders, the fecundity of affected men was lowered more than that of affected women, a finding that had been noted in previous studies.[14,40] This sex-specific effect suggests that psychiatric morbidity impairs interest or ability to find suitable mating partners or inhibits biological fertility to a greater extent in men. Several hypotheses attempt to explain why. Evolutionary theory suggests that male species have the potential for greater variance in reproductive success than female species.[41] This is assumed to result from their minimal investment being cheap compared with the minimal investment of women (ie, the cost of sperm production compared with 9 months of pregnancy). Because of this underlying difference in investment, it benefits women to be more selective in their choices of a mate. This can lead to mating systems often seen in mammals where "dominant" males have multiple mates, while others have none. Because of the greater variability in male fitness and the pressure on females to be selective, a genetic or environmental burden can have an exaggerated effect on a male's ability to find a mate. However, it should be noted that in this data set men had a lower variance in the number of children they had compared with women. Men also had fewer children on average than women, suggesting differences in reporting. A mother must be present at the birth of a child, but paternal uncertainty can exist even when a father is named. Another hypothesis is that fathers with a psychiatric disorder may be less likely to be recorded as the child's parent. Last, male fertility may be more susceptible than female fertility to sexual adverse effects of psychiatric treatment.

SCHIZOPHRENIA

Our findings for schizophrenia suggest a decrease in the fecundity of affected individuals similar to that found in a recent meta-analysis.[14] Our results suggest a strong selection pressure to remove genetic variants associated with schizophrenia from the population (**Table 3**). This is further evidence for the role of recent or de novo mutations in the genetic susceptibility to schizophrenia that have neither reached the frequency of nor existed long enough to be removed from the population.[12,42] This could act in conjunction with common variants' having too small an effect size to experience negative selection.[38] Our analysis also found a slight increase in the fecundity among sisters of affected individuals and a decrease in the fecundity among their brothers, although the increase in the fecundity of sisters disappears after correction for comorbidities. This trend in sibling fecundity is in agreement with several previous studies.[23,30,43,44] We have 2 potential hypotheses on why this might be the case. First, owing to the high heritability of schizophrenia and its

greater prevalence and severity among men, some brothers may have mild symptoms of schizophrenia but lack this diagnosis in our database. Second, schizophrenia may be the result of sexually antagonistic genes that are beneficial to female fitness at the expense of male fitness. This would be in agreement with the observations that schizophrenia affects men more severely than women and that the effect on fitness of shared genes in siblings is dependent on sibling sex. However, this may only be a partial explanation because the benefit to sisters does not compensate for the reduction in fitness seen in the affected individuals or brothers (Table 3).

AUTISM

Individuals with autism showed the greatest reduction in fecundity among all examined disorders. This was not unexpected because previous investigations have shown that few individuals with autism ever married or had children (eg, as demonstrated in the study by Larsen and Mouridsen[45]). The pattern of fecundity in affected individuals and siblings is similar to that of schizophrenia, with a slight increase in sisters' fecundity (nonsignificant herein) and a decrease in brothers' fecundity. As discussed for schizophrenia, this may be reflective of sexually antagonistic genes or undiagnosed symptoms in brothers. Therefore, we propose that rare highly deleterious variants and sexually antagonistic polymorphisms may contribute to the genetic disposition to autism. The similarity to schizophrenia is notable because it has been proposed that the autistic and psychotic spectrums reflect 2 extremes of social cognition.[46] It is unclear whether our results reflect a similar etiology or, as suggested by Crespi and Badcock,[46] they are opposite extremes of the same trait and come under the same stabilizing selection pressure. That both disorders show evidence for sexual antagonism supports the proposal by Crespi and Badcock that they are the result of sexual conflict.

BIPOLAR DISORDER

Unlike the other highly heritable disorders, bipolar disorder showed a low reduction in fecundity among affected individuals. This was accompanied by increased fecundity among sisters of affected individuals. However, owing to the large amount of comorbidity (see the "Methods" section), when this was corrected for, the fecundity of affected individuals increased to just below that of the general population. This agrees with the results of another study[34] of affective psychosis in Sweden but disagrees with other studies.[40,47,48] These studies have differed slightly in terms of diagnostic criteria, which may explain these discrepancies. It has been suggested that the introduction of lithium as a treatment for bipolar disorder has led to improved functioning and, as a result, greater fecundity in those populations where treatment is available.[49]

DEPRESSION

Notably, depression was an exception to the 5 other studied disorders. Female depressed individuals showed no difference in fecundity compared with the general population and had a slight increase in the fecundity after correction for comorbid disorders. Male depressed individuals showed a small decrease in the fecundity, although this too disappeared after correction for comorbidities. This contradicts the estimates of reduced fertility, especially in women, obtained from clinical samples of depressed individuals[50] and in a study[40] similar in design to ours using population registers in Denmark. The *ICD* classification used herein did not include identification of postnatal depression, and it is unclear to what extent this may lead to an increase in female fertility among individuals with depression. Furthermore, siblings of both sexes showed increased fecundity, and when this was taken into account, we found no selection acting against depression (Table 3). Rather, genes associated with depression seem to be maintained in the population by balancing selection because the cost to affected individuals is roughly equal to the benefit to their siblings. If this is the case, it would be the first strong evidence for balancing selection in a psychiatric disorder. The exact mechanism by which siblings benefit is beyond the scope of our analysis and is a line of future investigation. It has been proposed by Allen and Badcock[51] that depression may be adaptive in eliciting support from others. In parallel to the unique lack of negative influence on fecundity at the population level, depression stands out as a psychiatric disorder for which direct genetic associations have been most difficult to identify.[52-55] We propose that genetic studies in depression may benefit from the exploration of genetic and environmental dependencies that may contribute to balancing selection. An alternative explanation could be that environmental factors shared by siblings are associated with both an increased risk of depression and a higher fecundity.

ANOREXIA NERVOSA

Our analysis of anorexia showed a decrease in the fecundity among affected individuals but no difference in sibling fecundity. Our estimates of FRs in anorexia (0.54 in men and 0.81 in women) were less severe than a previous estimate from a clinical sample.[56] Our calculations suggest that anorexia is under weaker negative selection relative to schizophrenia and autism (Table 3).

SUBSTANCE ABUSE

Substance abuse was associated with reduced fecundity in affected individuals, but we found evidence for significantly increased fecundity in siblings of both sexes. Our findings suggest that this increased fecundity in siblings almost entirely accounts for the cost to affected individuals, with only a slight decrease (−3%) in the frequency of these individuals' genes predicted each generation. Considering that most drugs are a new environmental exposure when seen from an evolutionary perspective, it is possible that there has been insufficient time for selection to act on risk alleles. However, some evidence in the case of alcohol metabolism indicates that selection has affected different human populations differently.[57] Because alcohol abuse is the most frequent form of substance abuse in Sweden, we can as-

sume there has been sufficient time for some selection to have occurred.[58] It has also been suggested that substance abuse is associated with risk-taking behavior in both sexes, including sexual risk taking.[59]

SUMMARY AND RELEVANCE FOR PSYCHIATRIC GENETICS

The results of our analyses have several implications for future genetic studies. It seems likely that different evolutionary mechanisms underlie the persistence of the various psychiatric disorders. This in turn suggests that their genetic architecture may differ, so it is not surprising that the search for causal variants has proved more fruitful in some disorders than in others. More specifically, it seems that genetic variants conferring liability to schizophrenia, autism, and anorexia nervosa are under strong selection to be removed from the population. The continued high prevalence of schizophrenia and autism despite this strong negative selection, in combination with the aforementioned association with increased parental age, suggests that a high rate of de novo mutations may be maintaining these disorders in the population. The possibility of sexually antagonistic genes in schizophrenia and autism suggests that studies might benefit from male-only analysis, without women, who may be unaffected by risk alleles. Bipolar disorder did not seem to be under such strong negative selection and, after correcting for comorbidities, did not show sex-specific effects or changes in sibling fecundity.

Depression and, to a lesser extent, substance abuse, seems to be maintained by genes that are beneficial under some circumstances (ie, in siblings) but detrimental in others (ie, affected individuals). This suggests that gene-environment or gene-gene interactions have a large role in these disorders, for which some supporting evidence exists in depression.[60,61] This would decrease the power of studies comparing cases and controls, where many controls might also carry the genes that are "causal" for depression but not have the necessary genetic or environmental background risk factors to develop the disorder. Genes that interact with the environment may provide not only susceptibility to negative environments but also the ability to thrive in positive environments.[60] If the beneficial aspect of these genes is opposite to the disorder itself, rather than acting on a separate phenotype, then selecting high-functioning individuals as supercontrols might even increase the frequency of causal genetic variants in the controls. However, at this stage we have no evidence why siblings of individuals with depression or substance abuse would have increased fitness, and the observation could result from shared environmental factors uncorrected for in this analysis. Overall, a focus on case-only and exposed-only studies (eg, as in the study by Caspi et al[62]) might be more successful in disentangling the genetics of these disorders.

Submitted for Publication: October 5, 2011; final revision received January 16, 2012; accepted February 24, 2012.
Published Online: November 12, 2012. doi:10.1001/jamapsychiatry.2013.268

Author Affiliations: Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry (Mr Power and Dr Uher), and Department of Psychosis Studies (Drs MacCabe, McGuffin, and Lewis), King's College London, London, England; and Departments of Medical Epidemiology and Biostatistics (Drs Kyaga and Lichtenstein) and Public Health Sciences (Dr Svensson), Centre for Violence Prevention (Dr Långström), and Division of Public Health Epidemiology (Dr Svensson), Karolinska Institutet, Stockholm, and Institute of Neuroscience and Physiology, The Sahlgrenska Academy at Gothenburg University, Gothenburg (Dr Landen), Sweden.
Correspondence: Robert A. Power, BSc, Social Genetic and Developmental Centre, Institute of Psychiatry, King's College London, De Creispigny Park, London SE5 8AF, England (robert.r.power@kcl.ac.uk).
Financial Disclosure: Dr McGuffin has received consulting fees and honoraria from GlaxoSmithKline and Lundbeck.
Funding/Support: This study was supported by the Medical Research Council of the United Kingdom (Mr Power).
Additional Contributions: This study was made possible by previous work at the Karolinska Institutet using the Multi-Generation Register and the Swedish Hospital Discharge Register.

REFERENCES

1. Huxley J, Mayr E, Osmond H, Hoffer A. Schizophrenia as genetic morphism. *Nature*. 1964;204(495):220-221.
2. Keller MC, Miller G. Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *Behav Brain Sci*. 2006;29(4):385-452.
3. Uher R. The role of genetic variation in the causation of mental illness: an evolution-informed framework. *Mol Psychiatry*. 2009;14(12):1072-1082.
4. Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA. A de novo paradigm for mental retardation. *Nat Genet*. 2010;42(12):1109-1112.
5. Crow TJ. The 'big bang' theory of the origin of psychosis and the faculty of language. *Schizophr Res*. 2008;102(1-3):31-52.
6. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TFC, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature*. 2009;461(7265):747-753.
7. Reichenberg A, Gross R, Weiser M, Bresnahan M, Silverman J, Harlap S, Rabinowitz J, Shulman C, Malaspina D, Lubin G, Knobler HY, Davidson M, Susser E. Advancing paternal age and autism. *Arch Gen Psychiatry*. 2006;63(9):1026-1032.
8. Durkin MS, Maenner MJ, Newschaffer CJ, Lee LC, Cunniff CM, Daniels JL, Kirby RS, Leavitt L, Miller L, Zahorodny W, Schieve LA. Advanced parental age and the risk of autism spectrum disorder. *Am J Epidemiol*. 2008;168(11):1268-1276.
9. Malaspina D, Harlap S, Fennig S, Heiman D, Nahon D, Feldman D, Susser ES. Advancing paternal age and the risk of schizophrenia. *Arch Gen Psychiatry*. 2001;58(4):361-367.
10. Frans EM, Sandin S, Reichenberg A, Lichtenstein P, Långström N, Hultman CM. Advancing paternal age and bipolar disorder. *Arch Gen Psychiatry*. 2008;65(9):1034-1040.
11. Crow JF. Development. There's something curious about paternal-age effects. *Science*. 2003;301(5633):606-607.
12. Rees E, Moskvina V, Owen MJ, O'Donovan MC, Kirov G. De novo rates and selection of schizophrenia-associated copy number variants. *Biol Psychiatry*. 2011;70(12):1109-1114.
13. Kyaga S, Lichtenstein P, Boman M, Hultman C, Långström N, Landén M. Creativity and mental disorder: family study of 300,000 people with severe mental disorder. *Br J Psychiatry*. 2011;199(5):373-379.

14. Bundy H, Stahl D, MacCabe JH. A systematic review and meta-analysis of the fertility of patients with schizophrenia and their unaffected relatives. *Acta Psychiatr Scand*. 2011;123(2):98-106.
15. Multi-Generation Register 2009: a description of contents and quality. 2010. http://www.scb.se/statistik/_publikationer/BE9999_2009A01_BR_BE96BR1003.pdf. Accessed March 28, 2012.
16. *Swedish Hospital Discharge Register*. Stockholm, Sweden: Centre for Epidemiology, National Board of Health and Welfare; 2006.
17. Dalman Ch, Broms J, Cullberg J, Allebeck P. Young cases of schizophrenia identified in a national inpatient register—are the diagnoses valid? *Soc Psychiatry Psychiatr Epidemiol*. 2002;37(11):527-531.
18. Ekholm B, Ekholm A, Adolfsson R, Vares M, Osby U, Sedvall GC, Jönsson EG. Evaluation of diagnostic procedures in Swedish patients with schizophrenia and related psychoses. *Nord J Psychiatry*. 2005;59(6):457-464.
19. Bulik CM, Sullivan PF, Tozzi F, Furberg H, Lichtenstein P, Pedersen NL. Prevalence, heritability, and prospective risk factors for anorexia nervosa. *Arch Gen Psychiatry*. 2006;63(3):305-312.
20. Daniels JL, Forssen U, Hultman CM, Cnattingius S, Savitz DA, Feychting M, Sparen P. Parental psychiatric disorders associated with autism spectrum disorders in the offspring [published correction appears in *Pediatrics*. 2008;122(5):1162]. *Pediatrics*. 2008;121(5):e1357-e1362 http://pediatrics.aappublications.org/content/121/5/e1357.long. Accessed March 27, 2012.
21. Osby U, Brandt L, Correia N, Ekbom A, Sparén P. Excess mortality in bipolar and unipolar disorder in Sweden. *Arch Gen Psychiatry*. 2001;58(9):844-850.
22. Fazel S, Långström N, Hjern A, Grann M, Lichtenstein P. Schizophrenia, substance abuse, and violent crime. *JAMA*. 2009;301(19):2016-2023.
23. Svensson AC, Lichtenstein P, Sandin S, Hultman CM. Fertility of first-degree relatives of patients with schizophrenia: a three generation perspective. *Schizophr Res*. 2007;91(1-3):238-245.
24. Sellgren C, Landén M, Lichtenstein P, Hultman CM, Långström N. Validity of bipolar disorder hospital discharge diagnoses: file review and multiple register linkage in Sweden. *Acta Psychiatr Scand*. 2011;124(6):447-453.
25. Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R. Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Arch Gen Psychiatry*. 2003;60(7):709-717.
26. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis* . Hoboken, NJ: John Wiley & Sons; 2004:291-320.
27. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*. 1988;44(4):1049-1060.
28. Wicks S, Hjern A, Dalman C. Social risk or genetic liability for psychosis? A study of children born in Sweden and reared by adoptive parents. *Am J Psychiatry*. 2010;167(10):1240-1246.
29. Statistics Sweden. Educational attainment of the population 2010: Stockholm has highest level of education [press release]. 2010. http://www.scb.se/Pages/PressRelease____315969.aspx. Accessed March 28, 2012.
30. Haukka J, Suvisaari J, Lönnqvist J. Fertility of patients with schizophrenia, their siblings, and the general population: a cohort study from 1950 to 1959 in Finland. *Am J Psychiatry*. 2003;160(3):460-463.
31. *STATA Statistical Software* [computer program] . Release 12. College Station, TX: StataCorp LP; 2011.
32. Armstrong C. Thousands of women sterilised in Sweden without consent. *BMJ*. 1997;315(7108):563 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2127409/?tool=pubmed. Accessed March 27, 2012.
33. Tännsjö T. Compulsory sterilisation in Sweden. *Bioethics*. 1998;12(3):236-249.
34. MacCabe JH, Koupil I, Leon DA. Lifetime reproductive output over two generations in patients with psychosis and their unaffected siblings: the Uppsala 1915-1929 Birth Cohort Multigenerational Study. *Psychol Med*. 2009;39(10):1667-1676.
35. Ng CH. The stigma of mental illness in Asian cultures. *Aust N Z J Psychiatry*. 1997;31(3):382-390.
36. Shibre T, Negash A, Kullgren G, Kebede D, Alem A, Fekadu A, Fekadu D, Madhin G, Jacobsson L. Perception of stigma among family members of individuals with schizophrenia and major affective disorders in rural Ethiopia. *Soc Psychiatry Psychiatr Epidemiol*. 2001;36(6):299-303.
37. Lichtenstein P, Yip BH, Björk C, Pawitan Y, Cannon TD, Sullivan PF, Hultman CM. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009;373(9659):234-239.
38. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P, Ruderfer DM, McQuillin A, Morris DW, O'Dushlaine CT, Corvin A, Holmans PA, Macgregor S, Gurling H, Blackwood DHR, Corvin A, Craddock NJ, Gill M, Hultman CM, Kirov GK, Lichtenstein P, Muir WJ, Owen MJ, Pato CN, Scolnick EM, St Clair D, Craddock NJ, Holmans PA, Williams NM, Georgieva L, Nikolov I, Norton N, Williams H, Toncheva D, Milanova V, Hultman CM, Lichtenstein P, Thelander EF, Sullivan P, Kenny E, Quinn EM, Gill M, Corvin A, Choudhury K, Datta S, Pimm J, Thirumalai S, Puri V, Krasucki R, Lawrence J, Quested D, Bass N, Crombie C, Fraser G, Kuan SL, Walker N, Blackwood DHR, Muir WJ, McGhee KA, Pickard B, Malloy P, Maclean AW, Van Beck M, Wray NR, Macgregor S, Visscher PM, Pato MT, Medeiros H, Middleton F, Carvalho C, Morley C, Fanous A, Conti D, Knowles JA, Ferreira CP, Macedo A, Azevedo MH, Kirby AN, Ferreira MAR, Daly MJ, Chambert K, Kuruvilla F, Gabriel SB, Ardlie K, Moran JL, Daly MJ, Scolnick EM; International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460(7256):748-752.
39. Visscher PM, Goddard ME, Derks EM, Wray NR. Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses [published online June 14, 2011]. *Mol Psychiatry*. 2011.
40. Laursen TM, Munk-Olsen T. Reproductive patterns in psychotic patients. *Schizophr Res*. 2010;121(1-3):234-240.
41. Bateman AJ. Intra-sexual selection in *Drosophila*. *Heredity (Edinb)*. 1948;2(pt 3):349-368.
42. Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, Gogos JA, Karayiorgou M. Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat Genet*. 2011;43(9):864-868.
43. Weiser M, Reichenberg A, Werbeloff N, Halperin D, Kravitz E, Yoffe R, Davidson M. Increased number of offspring in first degree relatives of psychotic individuals: a partial explanation for the persistence of psychotic illnesses. *Acta Psychiatr Scand*. 2009;119(6):466-471.
44. Bassett AS, Bury A, Hodgkinson KA, Honer WG. Reproductive fitness in familial schizophrenia. *Schizophr Res*. 1996;21(3):151-160.
45. Larsen FW, Mouridsen SE. The outcome in children with childhood autism and Asperger syndrome originally diagnosed as psychotic. A 30-year follow-up study of subjects hospitalized as children. *Eur Child Adolesc Psychiatry*. 1997;6(4):181-190.
46. Crespi B, Badcock C. Psychosis and autism as diametrical disorders of the social brain. *Behav Brain Sci*. 2008;31(3):241-261.
47. Howard LM, Kumar C, Leese M, Thornicroft G. The general fertility rate in women with psychotic disorders. *Am J Psychiatry*. 2002;159(6):991-997.
48. Slater E, Hare EH, Price JS. Marriage and fertility of psychiatric patients compared with national data. *Soc Biol*. 1971;18:S60-S73.
49. Surja AAS, El-Mallakh RS. Fertility and childhood bipolar disorder. *Med Hypotheses*. 2007;69(3):587-589.
50. Williams KE, Marsh WK, Rasgon NL. Mood disorders and fertility in women: a critical review of the literature and implications for future research. *Hum Reprod Update*. 2007;13(6):607-616.
51. Allen NB, Badcock PBT. Darwinian models of depression: a review of evolutionary accounts of mood and mood disorders. *Prog Neuropsychopharmacol Biol Psychiatry*. 2006;30(5):815-826.
52. Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, Weale ME, Schosser A, Paredes UM, Rivera M, Craddock N, Owen MJ, Jones L, Jones I, Korszun A, Aitchison KJ, Shi J, Quinn JP, Mackenzie A, Vollenweider P, Waeber G, Heath S, Lathrop M, Muglia P, Barnes MR, Whittaker JC, Tozzi F, Holsboer F, Preisig M, Farmer AE, Breen G, Craig IW, McGuffin P. Genome-wide association study of major recurrent depression in the U.K. population. *Am J Psychiatry*. 2010;167(8):949-957.
53. Shi J, Potash JB, Knowles JA, Weissman MM, Coryell W, Scheftner WA, Lawson WB, DePaulo JR Jr, Gejman PV, Sanders AR, Johnson JK, Adams P, Chaudhury S, Jancic D, Evgrafov O, Zvinyatskovskiy A, Ertman N, Gladis M, Neimanas K, Goodell M, Hale N, Ney N, Verma R, Mirel D, Holmans P, Levinson DF. Genome-wide association study of recurrent early-onset major depressive disorder. *Mol Psychiatry*. 2011;16(2):193-201.
54. Sullivan PF, de Geus EJ, Willemsen G, James MR, Smit JH, Zandbelt T, Arolt V, Baune BT, Blackwood D, Cichon S, Coventry WL, Domschke K, Farmer A, Fava M, Gordon SD, He Q, Heath AC, Heutink P, Holsboer F, Hoogendijk WJ, Hottenga JJ, Hu Y, Kohli M, Lin D, Lucae S, Macintyre DJ, Maier W, McGhee KA, McGuffin P, Montgomery GW, Muir WJ, Nolen WA, Nöthen MM, Perlis RH, Pirlo K, Posthuma D, Rietschel M, Rizzu P, Schosser A, Smit AB, Smoller JW, Tzeng JY, van Dyck R, Verhage M, Zitman FG, Martin NG, Wray NR, Boomsma DI, Penninx BW. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry*. 2009;14(4):359-375.
55. Wray NR, Pergadia ML, Blackwood DH, Penninx BW, Gordon SD, Nyholt DR, Ripke S, MacIntyre DJ, McGhee KA, Maclean AW, Smit JH, Hottenga JJ, Willemsen G, Middeldorp CM, de Geus EJ, Lewis CM, McGuffin P, Hickie IB, van den Oord EJ, Liu JZ, Macgregor S, McEvoy BP, Byrne EM, Medland SE, Statham DJ, Henders AK, Heath AC, Montgomery GW, Martin NG, Boomsma DI, Madden PA, Sullivan PF. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Mol Psychiatry*. 2012;17(1):36-48.
56. Brinch M, Isager T, Tolstrup K. Anorexia nervosa and motherhood: reproduction pattern and mothering behavior of 50 women. *Acta Psychiatr Scand*. 1988;77(5):611-617.
57. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome [published corrections appear in *PLoS Biol*. 2007;5

(6):e147 and 2006;4(4):e154]. *PLoS Biol.* 2006;4(3):e72 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1382018/?tool=pubmed. Accessed March 28, 2012.
58. Grann M, Fazel S. Substance misuse and violent crime: Swedish population study. *BMJ.* 2004;328(7450):1233-1234.
59. Ramrakha S, Paul C, Bell M, Dickson N, Moffitt TE, Caspi A. The relationship between multiple sex partners and anxiety, depression and substance dependence disorders: a cohort study [abstract]. *J Sex Med.* 2011;8(suppl 3):258. doi:10.1111/j.1743-6109.2011.02325.x.
60. Belsky J, Pluess M. Beyond diathesis stress: differential susceptibility to environmental influences. *Psychol Bull.* 2009;135(6):885-908.
61. Uher R. The implications of gene-environment interactions in depression: will cause inform cure? *Mol Psychiatry.* 2008;13(12):1070-1078.
62. Caspi A, Hariri AR, Holmes A, Uher R, Moffitt TE. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry.* 2010;167(5):509-527.

RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART B
Neuropsychiatric Genetics

A Recessive Genetic Model and Runs of Homozygosity in Major Depressive Disorder

Robert A. Power,[1]* Matthew C. Keller,[2,3] Stephan Ripke,[4,5] Abdel Abdellaoui,[6] Naomi R. Wray,[7] Patrick F. Sullivan,[8,9,10,11] MDD PGC Working Group, and Gerome Breen[1,12,13]

[1]MRC Social Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, United Kingdom

[2]Department of Psychology and Neuroscience, University of Colorado at Boulder, Boulder, Colorado

[3]Institute for Behavioral Genetics, University of Colorado at Boulder, Boulder, Colorado

[4]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts

[5]Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts

[6]Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands

[7]The University of Queensland, Queensland Brain Institute, Brisbane, Australia

[8]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

[9]Department of Genetics, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[10]Department of Psychiatry, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[11]Department of Epidemiology, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[12]NIHR Biomedical Research Centre for Mental Health, South London and Maudsley NHS Foundation Trust, London, United Kingdom

[13]Institute of Psychiatry, King's College London, London, United Kingdom

Manuscript Received: 22 July 2013; Manuscript Accepted: 22 November 2013

The authors declare no conflicts of interest.

Grant sponsor: NIMH; Grant numbers: MH085520, MH080403, MH072802; Grant sponsor: German Federal Ministry of Education and Research; Grant number: 01GS08144 01GS08147; Grant sponsor: NIMH R01; Grant numbers: MH061686, MH059542, MH075131, MH059552, MH059541, MH060912; Grant sponsor: Broad Institute Center; Grant number: U54 RR020278; Grant sponsor: National Alliance for Research on Schizophrenia and Depression; Grant sponsor: BMBF Program Molecular Diagnostics: Validation of Biomarkers for Diagnosis and Outcome in Major Depression; Grant number: 01ES0811; Grant sponsor: Bavarian Ministry of Commerce, and the Federal Ministry of Education and Research (BMBF) in the framework of the National Genome Research Network; Grant numbers: FKZ 01GS0481, 01GS08145; Grant sponsor: The Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Twin Register (NTR); Grant sponsor: The Netherlands Organization for Scientific Research; Grant numbers: 904-61-090, 985-10-002, 904-61-193, 480-04-004, 400-05-717, 912-100-20, 916-76-125; Grant sponsor: Spinozapremie; Grant number: 56-464-14192; Grant sponsor: Geestkracht program; Grant number: 10-000-1002; Grant sponsor: The Center for Medical Systems Biology (NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure, VU University's Institutes for Health and Care Research and Neuroscience Campus Amsterdam, BIC/BioAssist/RK; Grant number: 2008.024; Grant sponsor: The European Science Foundation; Grant number: EU/QLRT-2001-01254; Grant sponsor: The European Community's Seventh Framework Program; Grant number: FP7/2007-2013; Grant sponsor: ENGAGE; Grant number: HEALTH-F4-2007-201413; Grant sponsor: The European Science Council; Grant number: ERC 230374; Grant sponsor: Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health (partial support); Grant sponsor: GAIN and the NIMH; Grant number: MH081802; Grant sponsor: Australian National Health and Medical Research Council; Grant numbers: 241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496675, 496739, 552485, 552498, 613602, 613608, 613674, 619667; Grant sponsor: The Australian Research Council; Grant numbers: FT0991360, FT0991022; Grant sponsor: The FP-5 GenomEUtwin Project; Grant number: QLG2-CT- 2002-01254; Grant sponsor: The US National Institutes of Health; Grant numbers: AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, MH66206, DA12854, DA019951; Grant sponsor: The Center for Inherited Disease Research (Baltimore, MD, USA); Grant sponsor: UK Medical Research Council and GlaxoSmithKline; Grant number: G0701420; Grant sponsor: The National Institute for Health Research (NIHR) Specialist Biomedical Research Centre for Mental Health at the South London; Grant sponsor: Maudsley NHS Foundation Trust; Grant sponsor: The Institute of Psychiatry, King's College London; Grant sponsor: The UK Medical Research Council; Grant number: G0000647; Grant sponsor: European Commission Framework; Grant number: LSHB-CT-2003-503428; Grant sponsor: National Institute of Mental Health; Grant number: N01MH90003.

*Correspondence to:
Robert A. Power, MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London SE5 8AF, United Kingdom.
E-mail: robert.r.power@kcl.ac.uk

Article first published online in Wiley Online Library (wileyonlinelibrary.com): 30 January 2014
DOI 10.1002/ajmg.b.32217

Genome-wide association studies (GWASs) of major depressive disorder (MDD) have yet to identify variants that surpass the threshold for genome-wide significance. A recent study reported that runs of homozygosity (ROH) are associated with schizophrenia, reflecting a novel genetic risk factor resulting from increased parental relatedness and recessive genetic effects. Here, we explore the possibility of such a recessive model in MDD. In a sample of 9,238 cases and 9,521 controls reported in a recent mega-analysis of 9 GWAS we perform an analysis of ROH and common variants under a recessive model. Since evidence for association with ROH could reflect a recessive mode of action at loci, we also conducted a genome-wide association analyses under a recessive model. The genome-wide association analysis using a recessive model found no significant associations. Our analysis of ROH suggested that there was significant heterogeneity of effect across studies in effect ($P = 0.001$), and it was associated with genotyping platform and country of origin. The results of the ROH analysis show that differences across studies can lead to conflicting systematic genome-wide differences between cases and controls that are unaccounted for by traditional covariates. They highlight the sensitivity of the ROH method to spurious associations, and the need to carefully control for potential confounds in such analyses. We found no strong evidence for a recessive model underlying MDD.

Key words: runs of homozygosity; recessive risk model; major depression; inbreeding

How to Cite this Article:
Power RA, Keller MC, Ripke S, Abdellaoui A, Wray NR, Sullivan PF, Breen G. 2014. A recessive genetic model and runs of homozygosity in major depressive disorder. Am J Med Genet Part B 165B:157–166.

INTRODUCTION

Major depressive disorder (MDD) is one of the leading burdens of disease in the world, with a lifetime prevalence of ~15% [Kessler et al., 2003; Hasin et al., 2005]. It has been found to be moderately heritable, from 31% to 42% [Sullivan et al., 2000], though with greater heritability in severe, recurrent forms of the disorder [McGuffin et al., 1996; Levinson, 2006]. A recent mega-analysis of nine genome-wide association studies found no significant associations with individual genetic variants (Psychiatric Genomics Consortium MDD Working Group, 2012), compared to ~5 genome-wide significant associations in similar sized studies of other psychiatric disorders [Ripke et al., 2011; Sklar et al., 2011]. These association studies are conducted under an additive model, while the true effects of some risk variants may be recessive, for which individuals with two copies of an allele are at greater risk than would be predicted from twice a single allele's effect. In a fully recessive model only those with two copies of the risk allele are at risk, though there is also the possibility of partial recessive effects. As selection acts to remove deleterious alleles with respect to overall fitness from the population, genetic risk variants that are recessive can escape selection longer. Inbreeding within families (e.g. consanguineous marriages) often exposes such recessive alleles due to an increased likelihood of alleles at each locus being identical by descent. Until recently studies of inbreeding were focused on families or communities in which inbreeding is expressed relative to the founder generation, which is assumed to be unrelated and where inbreeding information was determined from self-reports or knowledge of pedigrees (pedigree inbreeding). For example, Rudan et al. [2003] found a higher incidence of six complex genetic diseases/disorders including MDD among Croatian villages with higher levels of pedigree inbreeding [Rudan et al., 2003]. By using genome-wide genotype data it is also possible to estimate an individual's inbreeding from more distant common ancestors to provide evidence for whether a recessive genetic model is more appropriate for a disorder.

One method to analyse the effect of inbreeding from genome-wide genotype data is to identify segments of continuous homozygous SNPs, reflecting blocks of the genome that are identical by descent from a common ancestor. Runs of homozygosity (ROH) capture inbreeding effects that are due to common or rare causal variants better than a simple measure of excess number of homozygous SNPs across the genome, which tends to only capture the recessive effects of common variants [Keller et al., 2011]. An association between percentage of genome covered by ROH (F_{ROH}) and schizophrenia has been reported [Keller et al., 2012]. Due to the possibility of genetic overlap between MDD and schizophrenia [Schulze et al., 2012; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013], a similar association between F_{ROH} and MDD might be expected. However, MDD has a lower heritability ($h^2 \sim 0.37$, Sullivan et al., 2000) than schizophrenia ($h^2 \sim 0.81$, Sullivan et al., 2003), which should attenuate genetic relationships. Moreover, some authors have suggested that MDD may not be under negative selection and the casual genetic variants may be beneficial in some circumstances [Nesse, 1999; Watson and Andrews, 2002; Belsky and Pluess, 2009; Power et al., 2013]. Here we look at the association between MDD and SNPs in 9,238 cases and 9,521 controls across nine studies (Table 1) [Psychiatric Genomics Consortium MDD Working Group, 2012] under a recessive genetic model. We also use genome-wide estimates of inbreeding to look for a consistent difference between cases and controls across nine studies of MDD, in order to support a recessive model of the disorder.

MATERIAL AND METHODS

Sample

In this report, we analyzed individual data from the nine discovery samples [Ising et al., 2009; Sullivan et al., 2009; Lewis et al., 2010; Muglia et al., 2010; Rietschel et al., 2010; Shi et al., 2011; Shyn et al., 2011; Wray et al., 2012] of the PGC-MDD that together comprise 9,238 cases and 9,521 controls. Full sample details are given in the supplementary materials of the original analysis (Psychiatric Genomics Consortium MDD Working Group, 2012), and are outlined in Table 1. All subjects were of European

TABLE 1. Summary Statistics of Each of the Nine Studies and Results of Runs of Homozygosity Analysis

Study	N cases	N controls	Platform	Recurrence (%)	Ascertainment	Country	Number of SNPs[a]	Mean F_{ROH} (%)[b]	SD of Mean F_{ROH} (%)[b]	Traditional F estimate (%)[c]	Mean size of ROH (kb)	SD for Mean size of kb	Ref
RADIANT UK	1,625	1,588	Illumina 610k	97.1	Clinical	UK	263,728	0.10	0.41	0.1	3,294	1,373	Lewis et al. [2010]
MDD2000 OIMR-610k	433	751	Illumina 610k	57.6	Population	Australia	275,259	0.09	0.19	−0.01	3,376	1,049	Wray et al. [2012]
MDD2000 OIMR-317k	1,017	960	Illumina 317k[d]	51.4	Population	Australia	210,075	0.09	0.24	0.06	3,330	937	Wray et al. [2012]
Radiant + Bonn/Mannheim	935	1,290	Illumina 610k	74.2	Clinical	Germany	260,561	0.10	0.30	0.06	3,580	1,565	Rietschel et al. [2010]
GAIN	1,694	1,767	Perlegen 600k	50.7	Clinical/population	The Netherlands	199,749	0.12	0.36	0.15	3,780	1,840	Sullivan et al. [2009]
GenRED	1,030	1,253	Affymetrix 6.0	100	Volunteers	USA	253,826	0.08	0.18	0.09	3,238	763	Shi et al. [2011]
GSK	887	864	Illumina 550k	100	Clinical	Germany	281,029	0.16	0.67	0.14	3,732	1,711	Muglia et al. [2010]
MPIP	376	537	Illumina 317k	36.7	Clinical	Germany	215,393	0.11	0.30	0.04	3,533	1,424	Ising et al. [2009]
STAR*D	1,241	511	Affymetrix 5.0	74.3	Clinical	USA	155,428	0.10	0.40	0.17	3,362	783	Shyn et al. [2011]

[a]Genotyping was on either Illumina 317k or Illumina 370k SNPs, but only SNPs from Illumina 317k were used for imputation.
[b]After QC and pruning on linkage disequilibrium.
[c]F_{ROH} is the percentage of the genome covered by runs of homozygosity.
[d]This measure of *F* reflects an individual's observed number of homozygous loci compared to that expected under Hardy–Weinberg equilibrium.

ancestry (as determined from genome-wide genotypes). Cases were required to have diagnoses of DSM-IV lifetime MDD established using structured diagnostic instruments from direct interviews by trained interviewers (two studies required recurrent MDD and one recurrent, early-onset MDD) or clinician-administered DSM-IV checklists. Studies ascertained cases mostly from clinical sources, and controls were largely randomly selected from the population and screened for lifetime history of MDD.

Method of ROH Calling and Analysis

Genotyping was described in the supplementary materials in the original analysis (Psychiatric Genomics Consortium MDD Working Group, 2012). All samples were genotyped with single nucleotide polymorphism (SNP) arrays of greater than 200 K genome-wide SNPs, with analysis restricted to polymorphic SNP probes. In the original analysis, imputation to the CEU HapMap3 reference sample [Altshuler et al., 2010], 1,235,109 autosomal SNPs, was performed using Beagle 3.0.4 [Browning and Browning, 2009]. In order to perform an association analysis under a recessive model or call runs of homozygosity (ROH), imputed SNP dosage data was converted to discrete genotype calls, keeping those SNPs with a probability of at least 0.95. The use of imputed SNPs helped to increase similarity of genomic coverage across studies. SNPs with a missingness of >2% or minor allele frequency (MAF) <5% were removed, as were then individuals with missingness over 2%. Prior to analysis SNPs were pruned for LD within PLINK, removing any SNPs with an R^2 0.90 with any other SNP in a 50 SNP window. The use of imputed data in ROH has previously been shown to give similar results to those restricting to only genotyped SNPs [Keller et al., 2012]. The calling of ROH and percentage of genome covered by ROH per individual (F_{ROH}) were derived within PLINK [Purcell et al., 2007] following the same method found to optimally detect effects of autozygosity, as described in Howrigan et al. [2011]. In particular, we used a series of sliding windows across the genome to call ROH within each individual separately. The size of the windows was set to 65 consecutive SNPs, so any single SNP would be found in 65 different windows. If at least four (>5%) of these windows contained entirely homozygous SNPs, then the SNP in question could be included within a ROH. Within windows, one missing SNP was allowed. To avoid false positives, only ROH with a minimum of 65 consecutive SNPs covering 2.3 Mb were used when calculating F_{ROH}. In addition, the required minimum density in a ROH was set at 200 kb per SNP and the maximum gap between two consecutive homozygous SNPs was set at 500 kb. The estimate of the total genome captured was 2.77×10^9 bp. The analysis was performed by study, using F_{ROH} as a predictor of case-control status in a logistic regression. Percentage of SNPs missing, a SNP-by-SNP measure of homozygosity determined by PLINK's—het command, and the first five ancestry-informative principal components were used as covariates. The SNP-by-SNP measure of homozygosity was included to correct for differences in genomic-homozygosity levels unrelated to inbreeding, such as DNA quality or population ancestry. A mixed model was also examined combining all samples, using study as a random effect. This analysis was performed in STATA [StataCorp., 2011].

Genome-Wide Recessive Model

The genome-wide recessive model analysis used the autosomal dosage data converted to genotype calls as described above in the analysis of ROH. Analyses was performed in PLINK [Purcell et al., 2007], using the–recessive command. The first five ancestry-informative principal components were included as covariates. Analysis was restricted to autosomes. Each study was analysed separately and then a meta-analysis was performed for each SNP across studies (using fixed effect *P*-value in PLINK). As the risk allele is set as the minor allele by default, and this may differ by study for alleles at frequencies near 0.5, we used the minor allele in the analysis of imputed data from the whole sample as a reference. A *P*-value $<5 \times 10^{-8}$ was considered as genome-wide significant. For this significance cut-off, we had 90% power to detect a relative risk of 1.47 for the rare recessive genotype for SNPs with a MAF from 0.3 to 0.5. However, power decreased rapidly for those alleles with lower MAF, with 90% power to detect those with a relative risk of 1.81 and MAF of 0.2, or with a relative risk of 2.21 and MAF of 0.15. For those SNPs with lower MAF, power reduced rapidly for a recessive model. Calculations were performed using CaTS Power Calculator [Skol et al., 2006].

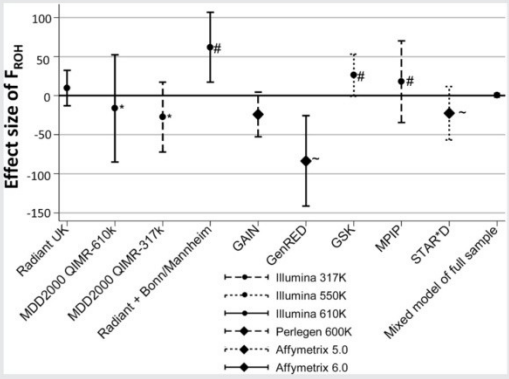


FIG. 1. Beta coefficient from logistic regression of F_{ROH} predicting MDD per study (with 95% CI, accounting for covariates). Positive effects suggest that ROHs are a risk factor for MDD. Note that though no combined effect in the mixed model of the full sample, in Illumina studies increased F_{ROH} was a risk factor ($P = 0.02$) and in non-Illumina studies it was protective ($P = 0.007$). Note also the consistency in studies from the same country: *Australian studies; ~American studies; #German studies.

RESULTS

Across all samples the average percentage of the genome covered by ROH (F_{ROH}) was 0.11% (95% CI 0.102–0.112; Table 1), similar albeit slightly lower than average F_{ROH} (0.15%) reported in an earlier report using the same parameters [Keller et al., 2012]. In our mixed model analysis across all samples with study as a random effect, we found no significant effect of F_{ROH} on MDD status. However, there was substantial heterogeneity in direction of effect across studies ($P = 0.001$, Fig. 1). Overall, four studies showed increased F_{ROH} in cases (one significantly, $P = 0.007$), while five studies showed increased F_{ROH} in controls (one where $P = 0.005$). Including further principal components (up to 20) and increasing ROH size (up to 170 SNPs) made no difference to the heterogeneity of the results.

To explore this apparent heterogeneity we examined two features of the included studies that might provide insight into the results. The first issue was potential poor matching of cases and controls, which we tested within the combined RADIANT German and Bonn-Mannheim sample. Here cases were recruited from both of these two studies, whereas the controls were collected and genotyped only as part of the Bonn–Mannheim study. However, excluding the RADIANT cases and restricting to only the matched Bonn–Mannheim study's cases and controls still resulted in a significant association with ROH ($P = 0.03$), as both sets of cases were found to have similar mean F_{ROH}. This tentatively suggested that the heterogeneity apparent across studies was not replicated within studies recruited from the same geographic region. Secondly, we were interested in the effects of genomic coverage on ROH. Cases and controls from the Queensland Institute of Medical Research (QIMR) were recruited as one sample but included as two independent cohorts, based on their genotyping platform (Illumina 317k and 610k chips). When we restricted the analysis to only those SNPs directly genotyped on both platforms (202,062 SNPs), we found that the mean F_{ROH} for the QIMR-610k sample reduced from 0.086% to 0.077%, compared to 0.078% for the QIMR-317k sample. This implies, as expected, that genotyping platform and genome coverage were influencing estimates of mean F_{ROH}.

To better understand potential sources of heterogeneity in our findings, we used a linear meta-regression with a study's effect size as the outcome and features of the studies individually analysed as predictors. We tested percentage of cases with recurrent MDD, mean F_{ROH} and genome-wide homozygosity, country of recruitment, presence of copy number variant probes on the platform, and genotyping platform as potential predictors of direction of effect (see Supplementary Table 1). Genotyping platform was nominally associated with effect size ($P = 0.05$), and the direction of effect of F_{ROH} differed when studies on non-Illumina platforms were analysed separately. Within the three non-Illumina genotyped studies, increased inbreeding was protective against MDD ($P = 0.007$) while within the six Illumina genotyped studies inbreeding was a significant risk factor for MDD ($P = 0.02$). We noted that when more than one study was recruited from a country, the direction of effect for ROH was consistent across all studies recruited from the same country (see Fig. 1). However, this was only a significant predictor in the meta-regression when distinguishing between German and non-German studies ($P = 0.02$). This level of confounding from genotyping and country likely reduced our power to detect a true effect of ROH and MDD.

In the meta-analysis of a recessive model of association, 929, 138 SNPs were analysed, though not all appeared in all nine studies due to differences in genotyping coverage. The most significant recessive association was for rs13261582 on chromosome 8 in an intergenic region between *SNTB1* and *HAS2* (odds ratio of 2.0 for the minor A allele, MAF 0.22, $P = 2.58 \times 10^{-6}$). However this SNP was only present in two studies (GenRED and STAR*D), and so is only supported by a subset of the sample. It did not appear among the reported top SNPs from the primary analysis of this dataset under an additive genetic model in the full sample ($P = 0.12$). The results of the meta-analysis also showed a lower median *P*-value than expected by chance (λ_{GC} 0.97, see Supplementary Figs. 1 and 2 for Manhattan and QQ plots). The λ_{GC} was 1.03 for SNPs with MAF > 0.2 and 0.90 for SNPs with MAF < 0.2, implying smaller differences of genotype frequencies between cases and controls than expected by chance for SNPS with low MAF. This is possibly the result of the less accurate imputation of rare alleles, or a lack of power.

DISCUSSION

Our analyses show systematic differences in F_{ROH} between cases and controls that differ in direction across studies. There are several explanations for these results, mostly highlighting limitations of this analysis. Firstly, we found systematic differences in mean F_{ROH} between studies. This is not unexpected and likely reflects the density of the genomic coverage and the accuracy of imputation, since SNPs were restricted to those with high quality imputation. A similar level of variation in F_{ROH} was observed in the Psychiatric GWAS Consortium's analysis of F_{ROH} across 17 studies of schizophrenia [Keller et al., 2012], though they did not report any heterogeneity of effect as a function of genotyping platform or country of recruitment. It seems unlikely that the heterogeneity of effect in the present study could be the result of differing SNP inclusion on the platforms, because such an explanation would imply systematic differences between cases and controls in the probability of homozygosity across SNPs as a function of platform. More likely in our opinion is the possibility that factors related to ascertainment of cases and controls differed across studies and influenced overall homozygosity. Such factors could include changes in homozygosity levels due to length and quality of DNA storage, or differences in ascertainment of cases and controls across populations. It is noteworthy that the two out of nine studies that genotyped controls independently of cases (GenRED and STAR*D) both showed higher F_{ROH} in controls than cases. Further, studies appear to cluster by country of origin and direction of the effect of F_{ROH}. All three German studies had increased F_{ROH} in cases for example, while the two Australian and two US studies all showed increased F_{ROH} in controls. This may reflect confounding demographic factors specific to each country. These unknown confounders, such as urban/rural status or religion, that influence both distant inbreeding (F_{ROH}) and MDD could explain the differences in effects between studies. A recent analysis of ROH and MDD in a partially overlapping sample of the GAIN study analysed here found exactly that. Religion confounded of the association due to reduced levels of depression but increased inbreeding in within the religious population of the Netherlands [Abdellaoui et al., 2013]. Certainly the initial hypothesis of this study, that an association with inbreeding would reflect negative selection on MDD and an excess of recessive causal mutations, seems an implausible explanation for the observed heterogeneity as the evolutionary cost of MDD status seems unlikely to have differed greatly among the ancestors of those included in the present study. Any of these explanations for the results of the F_{ROH} analysis may give some insight into why the original mega-analysis of these nine studies did not lead to any replicable genome-wide significant findings.

Our results from the genome-wide association analysis of MDD also produced no evidence for a recessive model, failing to produce any genome-wide significant associations. It is possible that our underlying model of recessive effect is unsuitable for an outbred population. Here we looked at a recessive effect for the minor allele, but two alternate models may also have been viable: compound heterozygosity and overdominance. Compound heterozygosity is an additional risk in individuals carrying two recessive but non-identical alleles within a genetic locus, while overdominance is the increased risk of homozygosity of any allele compared to being heterozygous. However, our analysis of both would have been restricted by low power and the use of biallelic markers, and were, therefore, not performed. Both the GWAS and ROH analyses suggest though that there is no underlying recessive model of MDD, at least not of large effect. Such an association was previously reported for schizophrenia in a similarly sized sample [Keller et al., 2012], showing an increase of risk for schizophrenia by 17% for every additional percentage of the genome covered by ROH and was taken as evidence for historical selection against schizophrenia risk variants. The lack of a similar association here adds molecular evidence to that from epidemiological studies suggesting MDD has little impact on reproductive fitness compared to other psychiatric disorders, and so is under substantially less negative selection [e.g., Power et al., 2013].

These results highlight that the analysis of F_{ROH} appears to be sensitive to systematic differences between studies that are ostensibly unrelated to MDD status, potentially give rise to either false positive or false negative results. This suggests there are genome-wide differences in homozygosity and/or inbreeding between populations that are not corrected for by methods such as ancestry-informative principal components. We recommend the use of large combined samples in the analysis of F_{ROH} as a predictor of traits and disorders, due to the high risk of spurious associations within one study. Preferably such analyses should be done with access to data on potential social and demographic confounders. One possible further improvement might be the development of novel methods for analysing ROH, particularly in imputed genotype data where probability for homozygosity across a region is available. As similar heterogeneity across studies was not seen in other analyses of ROH within consortia [Keller et al., 2012; McQuillan et al., 2012], the significant heterogeneity in our results suggest that MDD is particularly sensitive to differing demographics in the ascertainment of cases and controls and this may present a problem to genome-wide polygenic approaches such as ROH. Certainly no strong evidence for a recessive model was apparent, supporting the view of MDD being under weaker negative selection than other psychiatric disorders.

ACKNOWLEDGMENTS

Collaborators:

Michael R	Barnes	GlaxoSmithKline	UK
Thomas	Bettecken	Max Planck Institute of Psychiatry	Germany
Elisabeth B	Binder	Max Planck Institute of Psychiatry	Germany
Douglas HR	Blackwood	University of Edinburgh	UK
Dorret I	Boomsma	VU University, Amsterdam	The Netherlands
Gerome	Breen	Institute of Psychiatry, King's College London	UK
René	Breuer	Central Inst Mental Health, University of Heidelberg	Germany
Enda M	Byrne	Queensland Institute of Medical Research	Australia
Sven	Cichon	University of Bonn	Germany
William H	Coryell	University of Iowa	USA
Nick	Craddock	Cardiff University	UK
Ian W	Craig	Institute of Psychiatry, King's College London	UK
Darina	Czamara	Max Planck Institute of Psychiatry	Germany
Eco J	De Geus	VU University, Amsterdam	The Netherlands
Franziska	Degenhardt	University of Bonn	Germany
Anne E	Farmer	Institute of Psychiatry, King's College London	UK
Josef	Frank	Central Inst Mental Health, University of Heidelberg	Germany
Scott D	Gordon	Queensland Institute of Medical Research	Australia
Magdalena	Gross	University of Bonn	Germany
Steven P	Hamilton	University of California, San Francisco	USA
Andrew C	Heath	Washington University, St Louis	USA
Anjali K	Henders	Queensland Institute of Medical Research	Australia
Stefan	Herms	University of Bonn	Germany
Ian B	Hickie	University of Sydney, Sydney	Australia
Susanne	Hoefels	University of Bonn	Germany
Florian	Holsboer	Max Planck Institute of Psychiatry	Germany
Witte J	Hoogendijk	Erasmus Medical Center	The Netherlands
Jouke Jan	Hottenga	VU University, Amsterdam	The Netherlands
Marcus	Ising	Max Planck Institute of Psychiatry	Germany
Ian	Jones	Cardiff University	UK
Lisa	Jones	University of Birmingham	UK
Tzeng	Jung-Ying	North Carolina State University	USA
James A	Knowles	University of Southern California	USA
Martin A	Kohli	Max Planck Institute of Psychiatry	Germany
Ania	Korszun	Queen Mary University of London	UK
William B	Lawson	Howard University	USA
Douglas F	Levinson	Stanford University	USA
Cathryn M	Lewis	Institute of Psychiatry, King's College London	UK
Danyu	Lin	University of North Carolina	USA
Susanne	Lucae	Max Planck Institute of Psychiatry	Germany
Donald	MacIntyre	University of Edinburgh	UK
Pamela AF	Madden	Washington University, St Louis	USA
Wolfgang	Maier	University of Bonn	Germany
Nicholas G	Martin	Queensland Institute of Medical Research	Australia
Manuel	Mattheisen	University of Bonn	Germany
Patrick J	McGrath	Columbia University	USA
Peter	McGuffin	Institute of Psychiatry, King's College London	UK
Andrew	McIntosh	University of Edinburgh	UK
Alan	McLean	University of Edinburgh	UK
Christel M	Middeldorp	VU University, Amsterdam	The Netherlands
Lefkos	Middleton	Imperial College	UK
Grant M	Montgomery	Queensland Institute of Medical Research	Australia
Pierandrea	Muglia	GlaxoSmithKline	Italy
Bertram	Müller-Myhsok	Max Planck Institute of Psychiatry	Germany
Benjamin	Neale	Harvard University/Broad Institute	USA
Markus M	Noethen	University of Bonn	Germany
Willem A	Nolen	Groningen University Medical Center	The Netherlands
Dale R	Nyholt	Queensland Institute of Medical Research	Australia

Brenda P	Penninx	VU University Medical Center, Amsterdam	The Netherlands
Michele L	Pergadia	Washington University, St Louis	USA
James B	Potash	University of Iowa	USA
Marcella	Rietschel	Central Inst Mental Health, University of Heidelberg	Germany
Stephan	Ripke	Harvard University/Broad Institute	USA
William A	Scheftner	Rush University Medical Center	USA
Thomas G	Schulze	University of Goettingen	USA
Jianxin	Shi	National Cancer Institute	USA
Stanley I	Shyn	University of Washington	USA
Susan L	Slager	Mayo Clinic	USA
Johannes H	Smit	VU University Medical Center, Amsterdam	The Netherlands
Michael	Steffens	University of Bonn	Germany
Patrick F	Sullivan	University of North Carolina	USA
Federica	Tozzi	GlaxoSmithKline	Italy
Jens	Treutlein	Central Inst Mental Health, University of Heidelberg	Germany
Manfred	Uhr	Max Planck Institute of Psychiatry	Germany
Edwin JCG	van den Oord	Virginia Commonwealth University	USA
Gerard	VU University Medical Center, Amsterdam	The Netherlands	
Myrna M	Weissman	Columbia University	USA
Gonneke	Willemsen	VU University, Amsterdam	The Netherlands
Naomi R	Wray	The University of Queensland	Australia
Frans G	Zitman	Leiden University Medical Center, Leiden	The Netherlands

The PGC was funded by NIMH Grants MH085520 (lead PI PFS) and MH080403. The Bonn/Mannheim (BoMa) GWAS was supported by the German Federal Ministry of Education and Research, within the context of the National Genome Research Network 2 (NGFN-2), the National Genome Research Network plus (NGFNplus) and the Integrated Genome Research Network (IG) MooDS (Grant 01GS08144 to S. Cichon and M.M. No¨ethen, and Grant 01GS08147 to M. Rietschel). The GenRED GWAS project was supported by NIMH R01 Grants MH061686 (D.F. Levinson), MH059542 (W.H. Coryell), MH075131 (W.B. Lawson), MH059552 (J.B. Potash), MH059541 (W.A. Scheftner) and MH060912 (M.M. Weissman). We acknowledge the contributions of Dr. George S. Zubenko and Dr. Wendy N. Zubenko, Department of Psychiatry, University of Pittsburgh School of Medicine, to the GenRED I project. The NIMH Cell Repository at Rutgers University and the NIMH Center for Collaborative Genetic Studies on Mental Disorders made essential contributions to this project. Genotyping was carried out by the Broad Institute Center for Genotyping and Analysis with support from Grant U54 RR020278 (which partially subsidized the genotyping of the GenRED cases). Collection and quality control analyses of the control data set were supported by grants from NIMH and the National Alliance for Research on Schizophrenia and Depression. We are grateful to Knowledge Networks (Menlo Park, CA, USA) for assistance in collecting the control data set. We express our profound appreciation to the families who participated in this project, and to the many clinicians who facilitated the referral of participants to the study. Max Planck Institute of Psychiatry MARS study was supported by the BMBF Program Molecular Diagnostics: Validation of Biomarkers for Diagnosis and Outcome in Major Depression (01ES0811). Genotyping was supported by the Bavarian Ministry of Commerce, and the Federal Ministry of Education and Research (BMBF) in the framework of the National Genome Research Network (NGFN2 and NGFN-Plus, FKZ 01GS0481 and 01GS08145). The Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Twin Register (NTR) contributed to GAIN-MDD and to MDD2000. Funding was from: the Netherlands Organization for Scientific Research (Mag W./Zon M.W. Grants 904-61-090, 985-10- 002, 904-61-193, 480-04-004, 400-05-717, 912-100-20; Spinozapremie 56-464-14192; Geestkracht program Grant 10-000-1002); the Center for Medical Systems Biology (NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure, VU University's Institutes for Health and Care Research and Neuroscience Campus Amsterdam, BIC/BioAssist/RK (2008.024); the European Science Foundation (EU/QLRT-2001-01254); the European Community's Seventh Framework Program (FP7/2007-2013); ENGAGE (HEALTH-F4-2007-201413); and the European Science Council (ERC, 230374). Genotyping was funded in part by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health, and analysis was supported by grants from GAIN and the NIMH (MH081802). CM Middeldorp was supported by the Netherlands Organization for Scientific Research (NOW-VENI grant 916-76-125). Funding for the QIMR samples was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496675, 496739, 552485, 552498, 613602, 613608, 613674, 619667), the Australian Research Council (FT0991360, FT0991022), the FP-5 GenomEUtwin Project (QLG2-CT- 2002-01254) and the US National Institutes of Health (AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, MH66206, DA12854, DA019951), and the Center for Inherited Disease Research (Baltimore, MD, USA). We thank the twins and their families registered at the Australian Twin Registry for their participation in the many studies that have contributed to this research. RADIANT was funded by: a joint grant from the UK Medical

Research Council and GlaxoSmithKline (G0701420); the National Institute for Health Research (NIHR) Specialist Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and the Institute of Psychiatry, King's College London; and the UK Medical Research Council (G0000647). This work was funded in part by the National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health at South London and Maudsley NHS Foundation Trust and [Institute of Psychiatry] King's College London. This article/paper/report presents independent research in part funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The GENDEP study was funded by a European Commission Framework 6 grant, EC Contract Ref.: LSHB-CT-2003-503428. Genotyping of STAR*D wassupported by an NIMH Grant to S.P. Hamilton (MH072802). STAR*D was funded by the National Institute of Mental Health (contract N01MH90003) to the University of Texas Southwestern Medical Center at Dallas (A.J. Rush, principal investigator). We would like to thank the numerous researchers within the Psychiatric GWAS Consortium for their contributions. We also thank the thousands of people with MDD who donated time and effort to make this research possible.

REFERENCES

Abdellaoui A, Hottenga JJ, Xiao X, Scheet P, Ehli EA, Davies GE, Hudziak JJ, Smit DJ, Bartels M, Willemsen G, Brooks A, Sullivan PF, Smit JH, de Geus EJ, Penninx BW, Boomsma DI. 2013. Association between autozygosity and major depression: Stratification due to religious assortment. Behav Genet 43(6):455–467.

Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Bonnen PE, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarrol SA, Nemesh J, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Gonzaga-Jauregui C, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Zhang Q, Ghori MJ, McGinnis R, McLaren W, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE. 2010. Integrating common and rare genetic variation in diverse human populations. Nature 467(7311):52–58.

Belsky J, Pluess M. 2009. Beyond diathesis stress: Differential susceptibility to environmental influences. Psychol Bull 135(6):885–908.

Browning BL, Browning SR. 2009. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet 84(2):210–223.

Cross-Disorder Group of the Psychiatric Genomics Consortium. 2013. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nature Genetics 45:984–994 doi: 10.1038/ng.2711.

Hasin DS, Goodwin RD, Stinson FS, Grant BF. 2005. Epidemiology of major depressive disorder—Results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry 62(10):1097–1106.

Howrigan DP, Simonson MA, Keller MC. 2011. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. BMC Genomics 12:460.

Ising M, Lucae S, Binder EB, Bettecken T, Uhr M, Ripke S, Kohli MA, Hennings JM, Horstmann S, Kloiber S, Menke A, Bondy B, Rupprecht R, Domschke K, Baune BT, Arolt V, Rush AJ, Holsboer F, Muller-Myhsok B. 2009. A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. Arch Gen Psychiatry 66(9):966–975.

Keller MC, Simonson MA, Ripke S, Neale BM, Gejman PV, Howrigan DP, Lee SH, Lencz T, Levinson DF, Sullivan PF. 2012. Runs of homozygosity implicate autozygosity as a schizophrenia risk factor. PLoS Genet 8(4): e1002656.

Keller MC, Visscher PM, Goddard ME. 2011. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics 189(1):237–249.

Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS. 2003. The epidemiology of major depressive disorder—Results from the National Comorbidity Survey Replication (NCS-R). JAMA 289(23):3095–3105.

Levinson DF. 2006. The genetics of depression: A review. Biol Psychiatry 60(2):84–92.

Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, Weale ME, Schosser A, Paredes UM, Rivera M, Craddock N, Owen MJ, Jones L, Jones I, Korszun A, Aitchison KJ, Shi JX, Quinn JP, MacKenzie A, Vollenweider P, Waeber G, Heath S, Lathrop M, Muglia P, Barnes MR, Whittaker JC, Tozzi F, Holsboer F, Preisig M, Farmer AE, Breen G, Craig IW, McGuffin P. 2010. Genome-wide association study of major recurrent depression in the UK population. Am J Psychiatry 167(8):949–957.

McGuffin P, Katz R, Watkins S, Rutherford J. 1996. A hospital-based twin register of the heritability of DSM-IV unipolar depression. Arch Gen Psychiatry 53(2):129–136.

McQuillan R, Eklund N, Pirastu N, Kuningas M, McEvoy BP, Esko T, Corre T, Davies G, Kaakinen M, Lyytikainen LP, Kristiansson K, Havulinna AS, Gogele M, Vitart V, Tenesa A, Aulchenko Y, Hayward C, Johansson A, Boban M, Ulivi S, Robino A, Boraska V, Igl W, Wild SH, Zgaga L, Amin N, Theodoratou E, Polasek O, Girotto G, Lopez LM, Sala C, Lahti J, Laatikainen T, Prokopenko I, Kals M, Viikari J, Yang J, Pouta A, Estrada K, Hofman A, Freimer N, Martin NG, Kahonen M, Milani L, Heliovaara M, Vartiainen E, Raikkonen K, Masciullo C, Starr JM, Hicks AA, Esposito L, Kolcic I, Farrington SM, Oostra B, Zemunik T, Campbell H, Kirin M, Pehlic M, Faletra F, Porteous D, Pistis G, Widen E, Salomaa V, Koskinen S, Fischer K, Lehtimaki T, Heath A, McCarthy MI, Rivadeneira F, Montgomery GW, Tiemeier H, Hartikainen AL, Madden PA, d'Adamo P, Hastie ND, Gyllensten U, Wright AF, van Duijn CM, Dunlop M, Rudan I, Gasparini P, Pramstaller PP, Deary IJ, Toniolo D, Eriksson JG, Jula A, Raitakari OT, Metspalu A, Perola M, Jarvelin MR, Uitterlinden A, Visscher PM, Wilson JF. 2012. Evidence of inbreeding depression on human height. PLoS Genet 8(7): e1002655.

Muglia P, Tozzi F, Galwey NW, Francks C, Upmanyu R, Kong XQ, Antoniades A, Domenici E, Perry J, Rothen S, Vandeleur CL, Mooser V, Waeber G, Vollenweider P, Preisig M, Lucae S, Muller-Myhsok B, Holsboer F, Middleton LT, Roses AD. 2010. Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. Mol Psychiatry 15(6):589–601.

Nesse RM. 1999. Proximate and evolutionary studies of anxiety, stress and depression: Synergy at the interface. Neurosci Biobehav Rev 23(7):895–903.

Power RA, Kyaga S, Uher R, Maccabe JH, Langstrom N, Landen M, McGuffin P, Lewis CM, Lichtenstein P, Svensson AC. 2013. Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. Arch Gen Psychiatry 70(1):22–30.

Psychiatric Genomics Consortium MDD Working Group. 2012. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 18(4):497–511.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81(3):559–575.

Rietschel M, Mattheisen M, Frank J, Treutlein J, Degenhardt F, Breuer R, Steffens M, Mier D, Esslinger C, Walter H, Kirsch P, Erk S, Schnell K, Herms S, Wichmann HE, Schreiber S, Jockel KH, Strohmaier J, Roeske D, Haenisch B, Gross M, Hoefels S, Lucae S, Binder EB, Wienker TF, Schulze TG, Schmal C, Zimmer A, Juraeva D, Brors B, Bettecken T, Meyer-Lindenberg A, Muller-Myhsok B, Maier W, Nothen MM, Cichon S. 2010. Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. Biol Psychiatry 68(6):578–585.

Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA, Lin DY, Duan J, Ophoff RA, Andreassen OA, Scolnick E, Cichon S, St Clair D, Corvin A, Gurling H, Werge T, Rujescu D, Blackwood DH, Pato CN, Malhotra AK, Purcell S, Dudbridge F, Neale BM, Rossin L, Visscher PM, Posthuma D, Ruderfer DM, Fanous A, Stefansson H, Steinberg S, Mowry BJ, Golimbet V, De Hert M, Jonsson EG, Bitter I, Pietilainen OP, Collier DA, Tosato S, Agartz I, Albus M, Alexander M, Amdur RL, Amin F, Bass N, Bergen SE, Black DW, Borglum AD, Brown MA, Bruggeman R, Buccola NG, Byerley WF, Cahn W, Cantor RM, Carr VJ, Catts SV, Choudhury K, Cloninger CR, Cormican P, Craddock N, Danoy PA, Datta S, de Haan L, Demontis D, Dikeos D, Djurovic S, Donnelly P, Donohoe G, Duong L, Dwyer S, Fink-Jensen A, Freedman R, Freimer NB, Friedl M, Georgieva L, Giegling I, Gill M, Glenthoj B, Godard S, Hamshere M, Hansen M, Hansen T, Hartmann AM, Henskens FA, Hougaard DM, Hultman CM, Ingason A, Jablensky AV, Jakobsen KD, Jay M, Jurgens G, Kahn RS, Keller MC, Kenis G, Kenny E, Kim Y, Kirov GK, Konnerth H, Konte B, Krabbendam L, Krasucki R, Lasseter VK, Laurent C, Lawrence J, Lencz T, Lerer FB, Liang KY, Lichtenstein P, Lieberman JA, Linszen DH, Lonnqvist J, Loughland CM, Maclean AW, Maher BS, Maier W, Mallet J, Malloy P, Mattheisen M, Mattingsdal M, McGhee KA, McGrath JJ, McIntosh A, McLean DE, McQuillin A, Melle I, Michie PT, Milanova V, Morris DW, Mors O, Mortensen PB, Moskvina V, Muglia P, Myin-Germeys I, Nertney DA, Nestadt G, Nielsen J, Nikolov I, Nordentoft M, Norton N, Nothen MM, O'Dushlaine CT, Olincy A, Olsen L, O'Neill FA, Orntoft TF, Owen MJ, Pantelis C, Papadimitriou G, Pato MT, Peltonen L, Petursson H, Pickard B, Pimm J, Pulver AE, Puri V, Quested D, Quinn EM, Rasmussen HB, Rethelyi JM, Ribble R, Rietschel M, Riley BP, Ruggeri M, Schall U, Schulze TG, Schwab SG, Scott RJ, Shi J, Sigurdsson E, Silverman JM, Spencer CC, Stefansson K, Strange A, Strengman E, Stroup TS, Suvisaari J, Terenius L, Thirumalai S, Thygesen JH, Timm S, Toncheva D, van den Oord E, van Os J, van Winkel R, Veldink J, Walsh D, Wang AG, Wiersma D, Wildenauer DB, Williams HJ, Williams NM, Wormley B, Zammit S, Sullivan PF, O'Donovan MC, Daly MJ, Gejman PV. 2011. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 43(10):969–976.

Rudan I, Rudan D, Campbell H, Carothers A, Wright A, Smolej-Narancic N, Janicijevic B, Jin L, Chakraborty R, Deka R, Rudan P. 2003. Inbreeding and risk of late onset complex disease. J Med Genet 40(12):925–932.

Schulze TG, Akula N, Breuer R, Steele J, Nalls MA, Singleton AB, Degenhardt FA, Nothen MM, Cichon S, Rietschel M, McMahon FJ. 2012. Molecular genetic overlap in bipolar disorder, schizophrenia, and major depressive disorder. World J Biol Psychiatry. [epub ahed of print]

Shi J, Potash JB, Knowles JA, Weissman MM, Coryell W, Scheftner WA, Lawson WB, DePaulo JR, Gejman PV, Sanders AR, Johnson JK, Adams P, Chaudhury S, Jancic D, Evgrafov O, Zvinyatskovskiy A, Ertman N, Gladis M, Neimanas K, Goodell M, Hale N, Ney N, Verma R, Mirel D, Holmans P, Levinson DF. 2011. Genome-wide association study of recurrent early-onset major depressive disorder. Mol Psychiatry 16(2): 193–201.

Shyn SI, Shi J, Kraft JB, Potash JB, Knowles JA, Weissman MM, Garriock HA, Yokoyama JS, McGrath PJ, Peters EJ, Scheftner WA, Coryell W, Lawson WB, Jancic D, Gejman PV, Sanders AR, Holmans P, Slager SL, Levinson DF, Hamilton SP. 2011. Novel loci for major depression identified by genome-wide association study of Sequenced Treatment Alternatives to Relieve Depression and meta-analysis of three studies. Mol Psychiatry 16(2):202–215.

Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N, Edenberg HJ, Nurnberger JI Jr, Rietschel M, Blackwood D, Corvin A, Flickinger M, Guan W, Mattingsdal M, McQuillin A, Kwan P, Wienker TF, Daly M, Dudbridge F, Holmans PA, Lin D, Burmeister M, Greenwood TA, Hamshere ML, Muglia P, Smith EN, Zandi PP, Nievergelt CM, McKinney R, Shilling PD, Schork NJ, Bloss CS, Foroud T, Koller DL, Gershon ES, Liu C, Badner JA, Scheftner WA, Lawson WB, Nwulia EA, Hipolito M, Coryell W, Rice J, Byerley W, McMahon FJ, Schulze TG, Berrettini W, Lohoff FW, Potash JB, Mahon PB, McInnis MG, Zollner S, Zhang P, Craig DW, Szelinger S, Barrett TB, Breuer R, Meier S, Strohmaier J, Witt SH, Tozzi F, Farmer A, McGuffin P, Strauss J, Xu W, Kennedy JL, Vincent JB, Matthews K, Day R, Ferreira MA, O'Dushlaine C, Perlis R, Raychaudhuri S, Ruderfer D, Hyoun PL, Smoller JW, Li J, Absher D, Thompson RC, Meng FG, Schatzberg AF, Bunney WE, Barchas JD, Jones EG, Watson SJ, Myers RM, Akil H, Boehnke M, Chambert K, Moran J, Scolnick E, Djurovic S, Melle I, Morken G, Gill M, Morris D, Quinn E, Muhleisen TW, Degenhardt FA, Mattheisen M, Schumacher J, Maier W, Steffens M, Propping P, Nothen MM, Anjorin A, Bass N, Gurling H, Kandaswamy R, Lawrence J, McGhee K, McIntosh A, McLean AW, Muir WJ, Pickard BS, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Jones IR, Kirov G, Moskvina V, Nikolov I, O'Donovan MC, Owen MJ, Collier DA, Elkin A, Williamson R, Young AH, Ferrier IN, Stefansson K, Stefansson H, Thornorgeirsson T, Steinberg S, Gustafsson O, Bergen SE, Nimgaonkar V, Hultman C, Landen M, Lichtenstein P, Sullivan P, Schalling M, Osby U, Backlund L, Frisen L, Langstrom N, Jamain S, Leboyer M, Etain B, Bellivier F, Petursson H, Sigur Sson E, Muller-Mysok B, Lucae S, Schwarz M, Schofield PR, Martin N, Montgomery GW, Lathrop M, Oskarsson H, Bauer M, Wright A, Mitchell PB, Hautzinger M, Reif A, Kelsoe JR, Purcell SM. 2011. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 43(10): 977–983.

Skol AD, Scott LJ, Abecasis GR, Boehnke M. 2006. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 38(2):209–213.

StataCorp. 2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.

Sullivan PF, de Geus EJC, Willemsen G, James MR, Smit JH, Zandbelt T, Arolt V, Baune BT, Blackwood D, Cichon S, Coventry WL, Domschke K, Farmer A, Fava M, Gordon SD, He Q, Heath AC, Heutink P, Holsboer F, Hoogendijk WJ, Hottenga JJ, Hu Y, Kohli M, Lin D, Lucae S, MacIntyre DJ, Maier W, McGhee KA, McGuffin P, Montgomery GW, Muir WJ, Nolen WA, Nothen MM, Perlis RH, Pirlo K, Posthuma D, Rietschel M, Rizzu P, Schosser A, Smit AB, Smoller JW, Tzeng JY, van Dyck R, Verhage M, Zitman FG, Martin NG, Wray NR, Boomsma DI, Penninx BWJH. 2009. Genome-wide association for major depressive disorder: A possible role for the presynaptic protein piccolo. Mol Psychiatry 14(4):359–375.

Sullivan PF, Kendler KS, Neale MC. 2003. Schizophrenia as a complex trait—Evidence from a meta-analysis of twin studies. Arch Gen Psychiatry 60(12):1187–1192.

Sullivan PF, Neale MC, Kendler KS. 2000. Genetic epidemiology of major depression: Review and meta-analysis. Am J Psychiatry 157(10):1552–1562.

Watson PJ, Andrews PW. 2002. Toward a revised evolutionary adaptationist analysis of depression: The social navigation hypothesis. J Affect Disord 72(1):1–14.

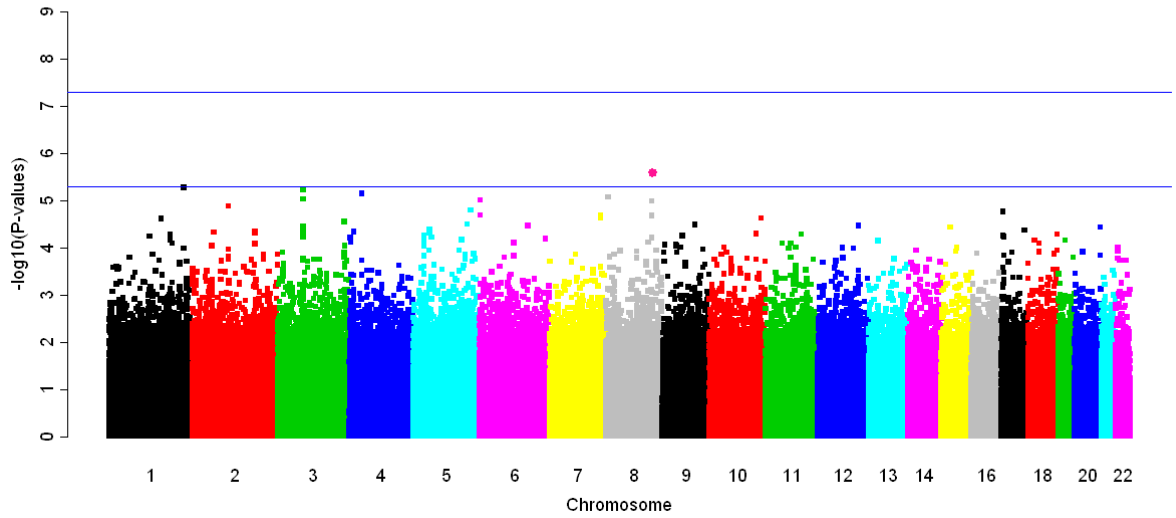
Wray NR, Pergadia ML, Blackwood DHR, Penninx BWJH, Gordon SD, Nyholt DR, Ripke S, MacIntyre DJ, McGhee KA, Maclean AW, Smit JH, Hottenga JJ, Willemsen G, Middeldorp CM, de Geus EJC, Lewis CM, McGuffin P, Hickie IB, van den Oord EJCG, Liu JZ, Macgregor S, McEvoy BP, Byrne EM, Medland SE, Statham DJ, Henders AK, Heath AC, Montgomery GW, Martin NG, Boomsma DI, Madden PAF, Sullivan PF. 2012. Genome-wide association study of major depressive disorder: New results, meta-analysis, and lessons learned. Mol Psychiatry 17(1):36–48.

SUPPORTING INFORMATION

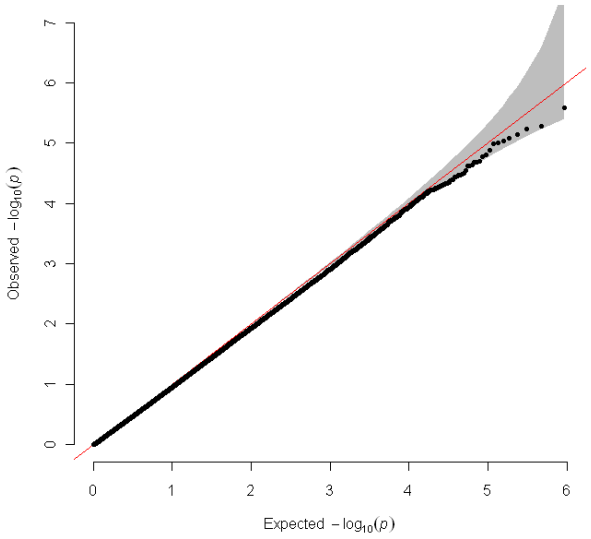
Additional supporting information may be found in the online version of this article at the publisher's web-site.

Supplementary materials

Supplementary Figure 1: Manhattan plot from p-values of genome-wide association analysis of major depression under a recessive model.



Supplementary Figure 2: QQ plot of p-values from genome-wide association analysis of major depression under a recessive model.



Supplementary Table 1:

Predictor	R^2	F_{ROH} increased in:	p-value
Percentage of recurrent cases	0.01	Controls	0.84
German study binary variable	0.58	Cases	0.02
US study binary variable	0.41	Controls	0.06
Australian study binary variable	0.04	Controls	0.59
Binary variable for if study was genotyped on Illumina	0.45	Cases	0.05
Binary variable for if copy number variant probes were available on the genotyping platform	0.01	Controls	0.79
Mean F_{ROH}	0.18	Cases	0.25
Mean genome-wide homozygosity	0.01	Controls	0.77

Journal of Psychiatric Research 47 (2013) 1032–1035

Contents lists available at SciVerse ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/psychires

The interaction between child maltreatment, adult stressful life events and the 5-*HTTLPR* in major depression

Robert A. Power [a,*], Lucy Lecky-Thompson [a], Helen L. Fisher [a], Sarah Cohen-Woods [a], Georgina M. Hosang [a], Rudolf Uher [a,b], Georgia Powell-Smith [a], Robert Keers [a,c], Maria Tropeano [a], Ania Korszun [c], Lisa Jones [d], Ian Jones [e], Michael J. Owen [e], Nick Craddock [e], Ian W. Craig [a], Anne E. Farmer [a], Peter McGuffin [a]

[a] MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, UK
[b] Department of Psychiatry, Dalhousie University, Halifax, NS, Canada
[c] Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Centre for Psychiatry, London, UK
[d] Department of Psychiatry, School of Clinical & Experimental Medicine, University of Birmingham, Birmingham, UK
[e] MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University, UK

ARTICLE INFO

Article history:
Received 30 December 2012
Received in revised form
26 March 2013
Accepted 27 March 2013

Keywords:
Mismatch hypothesis
Stressful life events
Child maltreatment
Major depression
Gene–environment interaction
5-*HTTLPR*

ABSTRACT

Both childhood maltreatment and adult stressful life events are established risk factors for the onset of depression in adulthood. However, the interaction between them can be viewed through two conflicting frameworks. Under a mismatch hypothesis stressful childhoods allow 'adaptive programming' for a stressful adulthood and so can be protective. Only when childhood and adulthood do not match is there a risk of behavioural problems. Alternatively, under the cumulative stress hypothesis we expect increased risk with each additional stressor. It has also been suggested that an individual's genetic background may determine the extent they undergo adaptive programming, and so which of these two hypotheses is relevant. In this study we test for an interaction between exposure to childhood maltreatment and adult stressful life events in a retrospective sample of 455 individuals, using major depression as the outcome. We also test whether this interaction differs by genotype at the 5-*HTTLPR*, a candidate for an individual's plasticity to adaptive programming.

Early maltreatment and stressful life events in adulthood interacted to produce increased risk for depression over each individually ($p = 0.055$). This supports the cumulative stress hypothesis over the mismatch hypothesis, at least with respect to severe environmental risk factors. This effect was not altered by 5-*HTTLPR* allele, suggesting there was no difference by genotype in adaptive programming to these events. We suggest that the apparent additional vulnerability to stressful events of those who have experienced maltreatment has clinical relevance, highlighting the importance of providing support beyond the immediate aftermath of maltreatment into adulthood.

1. Introduction

Child maltreatment (such as abuse or neglect) has long been associated with increased risk of psychopathology in adulthood, including depression (Edwards et al., 2003; Mazure, 1998; Windle et al., 1995), and presents a major social burden. Adult stressful life events (e.g. bereavement and divorce) also predict risk for the onset of depression (Brown and Harris, 1989; Monroe et al., 1999; Paykel et al., 1980) and triggering depressive episodes (Hosang et al., 2012). However, the interaction between these two types of adversity is far from simple. In a recent review of the topic, Nederhof and Schmidt (2012) described two conflicting models, the cumulative stress hypothesis and mismatch hypothesis, for explaining the interactions between early and adult stressors, and proposed a study design for differentiating them.

The cumulative stress hypothesis states that disease risk increases as adversity accumulates through life, the more traditional view for which there is considerable evidence e.g. (Brown et al., 2008; Kendler et al., 2004; McLaughlin et al., 2010). The mismatch hypothesis stipulates that individuals are primed to undergo "adaptive phenotypic programming" during development, in order to best match their adult phenotype with the predicted

* Corresponding author.
E-mail address: robert.r.power@kcl.ac.uk (R.A. Power).

0022-3956/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jpsychires.2013.03.017

environment (Frankenhuis and Del Giudice, 2012). In this case, a stressful childhood leads to developmental changes to prepare for a stressful adulthood. In contrast to the cumulative risk hypothesis, in the mismatch hypothesis individuals are more likely to experience disease if the early environment is different from the later adult environment; when the predicted programming and experienced environment are not in agreement. Here, one might imagine an individual adapted for low levels of stress responding poorly to adversity. Alternatively, an individual adapted for high levels of stress but encountering none may experience an apparent excess of anxiety or paranoia relative to their stress free environment. A recent review of animal models testing the mismatch hypothesis highlighted numerous examples in agreement with the model, but also many that were not (Schmidt, 2012).

One potentially confounding factor is that individuals have been predicted to differ genetically in their ability to experience adaptive programming and susceptibility to the environment (Belsky et al., 2009). Those who are genetically very susceptible to programming are more likely to follow the mismatch hypothesis (so called 'plastic' individuals), whereas those with no genetic susceptibility to programming do not undergo the initial adaptation and so the adverse effects of stress will accumulate. An individual's burden of 'plasticity polymorphisms' may therefore determine whether they respond according to the mismatch or the cumulative hypothesis, and may explain some of the conflicting results to date. One such genetic plasticity factor is a polymorphism within the promoter region of the serotonin transporter gene (5-*HTTLPR*), a candidate for plasticity suggested by Nederhof and Schmidt and previously found to interact with environmental risk factors for depression (Caspi et al., 2003; Eley et al., 2004; Fisher et al., 2013; Karg et al., 2011; Schmidt, 2012; Taylor et al., 2006; Uher et al., 2011).

Here we seek to follow Nederhof and Schmidt's (2012) proposed study design to test the cumulative risk model and the mismatch model in explaining major depression, including an attempt to incorporate an individual's genetic susceptibility to adaptive programming. We use reported levels of childhood maltreatment and stressful life events prior to interview/depressive episode as markers for early and adult stressors, respectively. As well as looking at their interaction in the whole sample, we also test stratifying by genotype at the 5-*HTTLPR* which, as mentioned, has been explored as a candidate for gene–environment interactions including within this sample (Fisher et al., 2013, 2012). To our knowledge, this will be the first human study incorporating genotypic data into a test of the mismatch hypothesis. A finding in support of either the cumulative stress or mismatch hypothesis has relevance to how individuals who have experienced childhood stress are likely to be affected by adult stress and risk of depression.

2. Methods

2.1. Participants

Individuals with recurrent unipolar depression and healthy controls were drawn from the Cardiff and London sites of the Depression Case–Control (DeCC) multi-centre study (Cohen-Woods et al., 2009). This study was approved by the local University and NHS Ethics Committees at each site and all participants provided written informed consent. Patients were identified through psychiatric clinics, hospitals, general medical practitioner surgeries, and media advertisements. Patients must have experienced at least two episodes of depression of at least moderate severity, separated by 2 or more months of remission, as defined by DSM-IV (American Psychiatric Association, 1994) and/or ICD-10 (World Health Organisation, 1993). All participants were aged 18 or over and had parents and grandparents of white European origin. Exclusion criteria were a history of mania or hypomania, mood-incongruent psychosis, and a first or second- degree relative with bipolar or psychotic disorder. Controls were recruited through UK general medical practices and excluded if they had a personal or first-degree relative with a history of any psychiatric disorder.

2.2. Measure of child maltreatment

Self-reported abuse or neglect during childhood, defined as until age 17, was recorded using the Childhood Trauma Questionnaire (CTQ) (Bernstein et al., 2003). The CTQ is a 28 item questionnaire widely used in clinical and general population samples and has good psychometric properties (Bernstein et al., 2003; Scher et al., 2004). Here we used a dichotomous variable to differentiate those who had experienced abuse or neglect from those who had not. This used a cut-off score of 8 or more for the sexual abuse, physical abuse or neglect subscales; a cut-off of 10 for emotional abuse; and a cut-off of 15 for the emotional neglect subscale. If an individual's scores exceeded any of these cut-offs, they were deemed to have experienced childhood maltreatment (Fisher et al., 2013).

2.3. Measure of stressful life events

The List of Threatening Experiences Questionnaire (LTE-Q) (Brugha et al., 1985) was used to record 11 stressful life events that occurred 6 months before cases' most severe episode of depression or 6 months prior to interview for controls (see Hosang et al., 2010). Each event was rated on a scale of zero (not present) to 3 (severe) and in this analysis, only severe events were counted. As in the measure of childhood maltreatment, the measure of stressful life events was dichotomized into those with any severe events and those with none (Fisher et al., 2012).

2.4. Genotyping

A 25 ml sample of whole blood was collected from cases at the time of interview and six cheek swabs were obtained from controls by mail. Polymerase chain reaction (PCR) was performed on the samples to amplify a 419 base-pair product for the l-allele (16-repeat) and a 375 base-pair product for the s-allele (14-repeat) of the 5-*HTTLPR* (Gelernter et al., 1997). The primer sequences were TGCCAGCACCTAACCCCTAATGT (forward) and GGACCGCAAGGTGG GCGGGA (reverse). The products were run on 2.5–3% agarose gel at 100 mV for 1 h. Genotyping was conducted blind to depression status and life events, and is described in previous publications (Fisher et al., 2013).

2.5. Analysis

In a logistic regression model we tested for each combination of stressors (childhood maltreatment only, adult stressors only, or both) against those who had no stressors, with depression as the outcome. We also tested whether an additive interaction between the dichotomous variables for childhood maltreatment and adult stressful life events was predictive of depression, using a binomial regression. Analyses were corrected for sex and age at assessment. Lastly, we tested to see if the direction of effect for these analyses differed depending on genotype at the 5-*HTTLPR*, using a dominant model for carriers of the short allele. We tested for normality of residuals and heteroscedasticity, and to account for the latter in some regressions we checked the results were consistent when robust standard errors were used. All analyses were carried out in STATA (StataCorp., 2011).

Table 1
Prevalence of environmental stressors and genotypes at 5-*HTTLPR* amongst cases and controls.

	N	Percentage female	Mean age at assessment (SE)	Mean age for adult SLE[a] (SE)	Mean child maltreatment score (SE)	Number reporting childhood maltreatment (%)	Mean adult SLE[a] score (SE)	Number reporting adult SLE[a] (%)	Number of 5-*HTTLPR* short allele carriers (%)
Cases	227	71.80%	45.4 (0.84)	37.9 (0.88)	46.0 (1.1)	184 (81%)	1.17 (0.08)	146 (65%)	157 (69%)
Controls	228	60.00%	47.2 (0.62)	47.2 (0.62)	33.0 (0 0.61)	73 (32%)	0.69 (0.07)	106 (46%)	146 (64%)

[a] SLE = stressful life events.

3. Results

The total number of individuals with phenotype and genotype data was 455, made up of 227 unipolar depression cases (71.8% female, mean age = 37.9, age range = 8–72) and 228 controls (60.1% female, mean age = 47.2, age range = 25–62). Of these, 257 individuals reported childhood maltreatment (71% cases) and 253 reported adult stressful life events prior to interview/worst episode of depression (58% cases) (see Table 1). The correlation between childhood maltreatment and adult stressful life events was small but significant ($r = 0.16$, $p = 0.0005$). The odds ratio for the prediction of adult stressful life events by childhood maltreatment in the whole sample was 1.74 ($p = 0.006$), but was only significant in cases (OR = 2.4, $p = 0.01$). The short allele of the 5-*HTTLLPR* had a frequency of 42.6%, and was in Hardy–Weinberg Equilibrium ($p = 0.66$). 5-*HTTLLPR* genotype was not associated with depression status.

Child maltreatment was associated with depression in those with and without stressful life events, but stressful life events were only linked to elevated rates of depression when combined with child maltreatment (see Table 2). Moreover, the odds were much greater when individuals reported exposure to both child and adult stressors than for child maltreatment alone, increasing from an OR of 5.1–12.4 (Table 2). There was tentative evidence of an interaction effect between childhood maltreatment and adult stressful life events (see Table 2, $p = 0.055$). The same effect size for this interaction was seen in both those who did and did not carry a copy of the short allele at the 5-*HTTLPR*, suggesting no difference in propensity towards either mismatch or cumulative hypothesis based on genotype. There was a much greater effect of child maltreatment on the risk of depression in those individuals carrying the short allele (from OR of 7.8–15.8), reflecting the gene–environment interaction of risk seen previously in this sample (Fisher et al., 2013).

4. Discussion

Our findings provide clear evidence against the mismatch hypothesis for major depression with respect to the environmental stressors examined here, and supports a view of cumulative risk from repeated adversity during an individual's lifespan. Furthermore, the results indicate an interaction effect, where adult stressful life events were only associated with depression in those individuals who had experienced childhood maltreatment. While we focused only on one candidate genetic polymorphism in the serotonin transporter gene (5-*HTTLPR*), we did not find evidence for an underlying genetic susceptibility to adaptive programming in response to childhood maltreatment and future stressful events.

The major limitation of this study is the measurements of early and late stressful events. Brief self-report questionnaires were utilised to retrospectively assess stressors which may have biased the information obtained. However, high concordances between the self-reported LTE-Q used here and more detailed measures of adult life events such as the Life Events and Difficulties Schedule have previously been reported (Hosang et al., 2010). Self-reporting of childhood maltreatment has also shown to agree with more objective measures in the form of hospital records (McGuffin et al., 1986). However the real issue is the severity of the measures of environment. Childhood maltreatment, as measured here, presents a very severe negative experience. It is likely that adaptive programming only occurs with respect to milder measures of childhood stress which are less damaging to an individual and allow them to recover. Unfortunately no such measures were available here, and we hope this study will promote the importance of detailed environmental phenotyping beyond the most severe risk factors. With regards to adult stressful life events we were limited to those events prior to the most severe depressive episode, and it is possible that we might have found different results if we had been able to use events prior to the initial onset of depression. Only for a very small proportion of this sample was the worst episode also the first episode (~7%), prohibiting such an analysis. While not ideal, the measure of events prior to the most severe episode still captures an individual's sensitivity to environmental stressors relative to controls and so should not be dismissed as uninformative. Lastly, this study focused solely on severe recurrent depression as an outcome and it is possible that the mismatch framework may be operating for other behavioural phenotypes.

Overall a better understanding of how risk factors interact for disorders such as depression is important in guiding the treatment and management of affected individuals. This is especially true where there is the possibility of adaptive programming, and so the possibility of environmental interventions. Here we have shown novel findings in support of the cumulative stress hypothesis explaining risk for depression over the mismatch hypothesis, and highlighted the increased risk of those individuals who have suffered childhood maltreatment. This also highlights the need to

Table 2
Odds ratios (OR) and *p*-values for depression when comparing those who have experienced no stressors ($n = 106$) to those who have experienced child maltreatment only ($n = 96$), adult stressful life events only ($n = 92$), or both ($n = 161$), and an interaction term for the two variables. Results also shown by 5-*HTTLPR* genotype in a dominant model for the short allele.

Adversity	Total sample		Long allele homozygotes ($n = 152$)		Short allele carriers ($n = 302$)	
	Or (95% CI)	*p*-value	Or (95% CI)	*p*-value	Or (95% CI)	*p*-value
Only adult	1.01 (0.50–2.04)	0.98	1.02 (0.33–3.18)	0.97	0.94 (0.37–2.36)	0.89
Only child	5.05 (2.56–9.94)	<0.001	2.7 (0.88–8.27)	0.08	7.3 (3.02–17.64)	<0.001
Both	12.41 (6.53–23.60)	<0.001	7.82 (2.62–23.23)	<0.001	15.75 (7.01–35.37)	<0.001
Interaction	0.92 (−0.02 to 1.88)	0.055	0.99 (−0.57 to 2.56)	0.21	0.93 (−0.28 to 2.14)	0.13

CI, confidence interval.

equip such individuals with the tools and support to deal with future stressful events beyond the immediate aftermath of such maltreatment, and into adulthood.

Role of funding

Funding for the DeCC study was provided by the UK Medical Research Council (MRC).

Contributors

Robert A. Power and Lucy Lecky-Thompson were involved in the writing of the manuscript and proposed the analysis. Helen L. Fisher, Sarah Cohen-Woods, Georgina M. Hosang, Rudolf Uher, Georgia Powell-Smith, Robert Keers and Maria Tropeano were involved in the analysis of both environmental and genetic data. Ania Korszun, Lisa Jones, Ian Jones, Mike Owen, Nick Craddock, Ian W. Craig, Anne E. Farmer and Peter McGuffin were involved in the overall study design, selection of measures, and recruitment of individuals.

Conflicts of interest

Anne Farmer and Peter McGuffin have received consultancy fees and honoraria for participating in expert panels from pharmaceutical companies including Lundbeck and GlaxoSmithKline. All other authors declare that they have no conflicts of interest. Georgina Hosang has provided a sponsored talk for Bristol-Myers Squibb.

Acknowledgements

None.

References

Association AP. Diagnostic and statistical manual of mental disorders 4th edition (DSM-IV). Washington DC: American Psychiatric Press; 1994.
Belsky J, Jonassaint C, Pluess M, Stanton M, Brummett B, Williams R. Vulnerability genes or plasticity genes? Molecular Psychiatry 2009;14:746–54.
Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. Child Abuse & Neglect 2003;27:169–90.
Brown GW, Craig TK, Harris TO. Parental maltreatment and proximal risk factors using the Childhood Experience of Care & Abuse (CECA) instrument: a life-course study of adult chronic depression – 5. Journal of Affective Disorders 2008;110:222–33.
Brown GW, Harris TO. Life events and illness. New York: Guildford Publications; 1989.
Brugha T, Bebbington P, Tennant C, Hurry J. The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. Psychological Medicine 1985;15:189–94.
Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 2003;301:386–9.
Cohen-Woods S, Gaysina D, Craddock N, Farmer A, Gray J, Gunasinghe C, et al. Depression Case Control (DeCC) Study fails to support involvement of the muscarinic acetylcholine receptor M2 (CHRM2) gene in recurrent major depressive disorder. Human Molecular Genetics 2009;18:1504–9.
Edwards V, Holden GW, Felitti VJ, Anda RF. Relationship between multiple forms of childhood maltreatment and adult mental health in community respondents: results from the adverse childhood experiences study. American Journal of Psychiatry 2003;160:1453–60.
Eley TC, Sugden K, Corsico A, Gregory AM, Sham P, McGuffin P, et al. Gene-environment interaction analysis of serotonin system markers with adolescent depression. Molecular Psychiatry 2004;9:908–15.
Fisher HL, Cohen-Woods S, Hosang GM, Korszun A, Owen M, Craddock N, et al. Interaction between specific forms of childhood maltreatment and the serotonin transporter gene (5-HTT) in recurrent depressive disorder. Journal of Affective Disorders 2013;145:136–41.
Fisher HL, Cohen-Woods S, Hosang GM, Uher R, Powell-Smith G, Keers R, et al. Stressful life events and the serotonin transporter gene (5-HTT) in recurrent clinical depression. Journal of Affective Disorders 2012;136:189–93.
Frankenhuis WE, Del Giudice M. When do adaptive developmental mechanisms yield maladaptive outcomes? Developmental Psychology 2012;48:628–42.
Gelernter J, Kranzler H, Cubells JF. Serotonin transporter protein (SLC6A4) allele and haplotype frequencies and linkage disequilibria in African- and European-American and Japanese populations and in alcohol-dependent subjects. Human Genetics 1997;101:243–6.
Hosang GM, Korszun A, Jones L, Jones I, Gray JM, Gunasinghe CM, et al. Adverse life event reporting and worst illness episodes in unipolar and bipolar affective disorders: measuring environmental risk for genetic research. Psychological Medicine 2010;40:1829–37.
Hosang GM, Korszun A, Jones L, Jones I, McGuffin P, Farmer AE. Life-event specificity: bipolar disorder compared with unipolar depression. British Journal of Psychiatry 2012;201:458–65.
Karg K, Burmeister M, Shedden K, Sen S. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. Archives of General Psychiatry 2011;68:444–54.
Kendler KS, Kuhn JW, Prescott CA. Childhood sexual abuse, stressful life events and risk for major depression in women. Psychological Medicine 2004;34:1475–82.
Mazure CM. Life stressors as risk factors in depression. Clinical Psychology: Science and Practice 1998;5:291–313.
McGuffin P, Katz R, Aldrich J. Past and present state examination: the assessment of 'lifetime ever' psychopathology. Psychological Medicine 1986;16:461–5.
McLaughlin KA, Conron KJ, Koenen KC, Gilman SE. Childhood adversity, adult stressful life events, and risk of past-year psychiatric disorder: a test of the stress sensitization hypothesis in a population-based sample of adults. Psychological Medicine 2010;40:1647–58.
Monroe SM, Rohde P, Seeley JR, Lewinsohn PM. Life events and depression in adolescence: relationship loss as a prospective risk factor for first onset of major depressive disorder. Journal of Abnormal Psychology 1999;108:606–14.
Nederhof E, Schmidt MV. Mismatch or cumulative stress: toward an integrated hypothesis of programming effects. Physiology & Behavior 2012;106:691–700.
Organisation, W. H. The ICD-10 classification of mental and behavioural disorders. Diagnostic criteria for research. Geneva: Switzerland:World Health Organisation; 1993.
Paykel ES, Emms EM, Fletcher J, Rassaby ES. Life events and social support in puerperal depression. British Journal of Psychiatry 1980;136:339–46.
Scher CD, Forde DR, McQuaid JR, Stein MB. Prevalence and demographic correlates of childhood maltreatment in an adult community sample. Child Abuse & Neglect 2004;28:167–80.
Schmidt MV. Animal models for depression and the mismatch hypothesis of disease. Psychoneuroendocrinology 2012;36:330–8.
StataCorp. Stata statistical software: release 12. College Station, TX: StataCorp LP; 2011.
Taylor SE, Way BM, Welch WT, Hilmert CJ, Lehman BJ, Eisenberger NI. Early family environment, current adversity, the serotonin transporter promoter polymorphism, and depressive symptomatology. Biological Psychiatry 2006;60:671–6.
Uher R, Caspi A, Houts R, Sugden K, Williams B, Poulton R, et al. Serotonin transporter gene moderates childhood maltreatment's effects on persistent but not single-episode depression: replications and implications for resolving inconsistent results. Journal of Affective Disorders 2011;135:56–65.
Windle M, Windle RC, Scheidt DM, Miller GB. Physical and sexual abuse and associated mental disorders among alcoholic inpatients. American Journal of Psychiatry 1995;152:1322–8.

RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART B
Neuropsychiatric Genetics

Genome-Wide Association Analysis Accounting for Environmental Factors Through Propensity-Score Matching: Application to Stressful Live Events in Major Depressive Disorder

Robert A. Power,[1]* Sarah Cohen-Woods,[1,2] Mandy Y. Ng,[1] Amy W. Butler,[1,3] Nick Craddock,[4] Ania Korszun,[5] Lisa Jones,[6] Ian Jones,[4] Michael Gill,[7] John P. Rice,[8] Wolfgang Maier,[9] Astrid Zobel,[9] Ole Mors,[10] Anna Placentino,[11] Marcella Rietschel,[12] Katherine J. Aitchison,[1] Federica Tozzi,[13] Pierandrea Muglia,[13,14,15] Gerome Breen,[1,16] Anne E. Farmer,[1] Peter McGuffin,[1] Cathryn M. Lewis,[1] and Rudolf Uher[1,17]

[1]MRC Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, United Kingdom

[2]Discipline of Psychiatry, School of Medicine, University of Adelaide, Adelaide, SA 5005, Australia

[3]Department of Psychiatry, University of Hong Kong, Hong Kong, Special Administrative Region, China

[4]MRC Centre for Neuropsychiatric Genetics and Genomics, Neuroscience and Mental Health Research Institute, Cardiff University, Cardiff, United Kingdom

[5]Barts and The London Medical School, Queen Mary University of London, London, United Kingdom

[6]Department of Psychiatry, School of Clinical & Experimental Medicine, University of Birmingham, Birmingham, United Kingdom

[7]Department of Psychiatry, Trinity Centre for Health Science, Dublin, Ireland

[8]Department of Psychiatry, Washington University, St. Louis, Missouri

[9]Department of Psychiatry, University of Bonn, Bonn, Germany

[10]Centre for Psychiatric Research, Aarhus University Hospital, Risskov, Denmark

[11]UOP 23-Department of Mental Health, Spedali Civili Hospital and Biological Psychiatry Unit-Centro San Giovanni di Dio, FBF, Brescia, Italy

[12]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

[13]GlaxoSmithKline Research & Development, Verona, Italy

[14]Department of Psychiatry, University of Toronto, Toronto, Canada

[15]NeuroSearch A/S, Ballerup, Denmark

Conflicts of interest: Aitchison, Farmer, and McGuffin have received consultancy fees and honoraria for participating in expert panels for pharmaceutical companies including GlaxoSmithKline. Aitchison's declares interests through Advisory Boards for Johnson & Johnson, Lundbeck, Roche Diagnostics, and Bristol-Myers Squibb; membership of Bristol-Myers Squibb UK Steering group 2003 to present; consultancy work for Roche Diagnostics, Johnson & Johnson Pharmaceutical Research and Development, Lundbeck, and Bristol-Myers Squibb Pharmaceuticals Limited; grants awarded by Johnson & Johnson Pharmaceutical Research & Development, Bristol-Myers Squibb Pharmaceuticals Limited, and E Merck Pharmaceuticals. Tozzi and Muglia were employees of GlaxoSmithKline when the research was performed. All other authors (Power, Butler, Ng, Cohen-Woods, Craddock, Korszun, Jones I, Jones L, Gill, Rice, Maier, Zobel, Mors, Placentino, Rietschel, Breen, Craig, Lewis, and Uher) declare no conflicts of interest.

Cathryn M. Lewis and Rudolf Uher contributed equally to this work.

Grant sponsor: Medical Research Council, UK (RADIANT); Grant sponsor: GlaxoSmithKline; Grant number: G0701420; Grant sponsor: NIHR Biomedical Research Centre for Mental Health at the South London; Grant sponsor: Maudsley NHS Foundation Trust; Grant sponsor: Institute of Psychiatry, King's College London.

*Correspondence to:
Robert A. Power, MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, London SE5 8AF, United Kingdom.
E-mail: robert.r.power@kcl.ac.uk

Article first published online in Wiley Online Library (wileyonlinelibrary.com): 15 July 2013
DOI 10.1002/ajmg.b.32180

[16]NIHR Biomedical Research Centre for Mental Health, South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, King's College London, London, United Kingdom

[17]Department of Psychiatry, Dalhousie University, Halifax, NS, Canada

Manuscript Received: 4 December 2012; Manuscript Accepted: 5 June 2013

How to Cite this Article:
Power RA, Cohen-Woods S, Ng MY, Butler AW, Craddock N, Korszun A, Jones L, Jones I, Gill M, Rice JP, Maier W, Zobel A, Mors O, Placentino A, Rietschel M, Aitchison KJ, Tozzi F, Muglia P, Breen G, Farmer AE, McGuffin P, Lewis CM, Uher R. 2013. Genome-Wide Association Analysis Accounting for Environmental Factors Through Propensity-Score Matching: Application to Stressful Live Events in Major Depressive Disorder. Am J Med Genet Part B 162B:521–529.

Stressful life events are an established trigger for depression and may contribute to the heterogeneity within genome-wide association analyses. With depression cases showing an excess of exposure to stressful events compared to controls, there is difficulty in distinguishing between "true" cases and a "normal" response to a stressful environment. This potential contamination of cases, and that from genetically at risk controls that have not yet experienced environmental triggers for onset, may reduce the power of studies to detect causal variants. In the RADIANT sample of 3,690 European individuals, we used propensity score matching to pair cases and controls on exposure to stressful life events. In 805 case–control pairs matched on stressful life event, we tested the influence of 457,670 common genetic variants on the propensity to depression under comparable level of adversity with a sign test. While this analysis produced no significant findings after genome-wide correction for multiple testing, we outline a novel methodology and perspective for providing environmental context in genetic studies. We recommend contextualizing depression by incorporating environmental exposure into genome-wide analyses as a complementary approach to testing gene-environment interactions. Possible explanations for negative findings include a lack of statistical power due to small sample size and conditional effects, resulting from the low rate of adequate matching. Our findings underscore the importance of collecting information on environmental risk factors in studies of depression and other complex phenotypes, so that sufficient sample sizes are available to investigate their effect in genome-wide association analysis.

Key words: depression; propensity score matching; stressful life events; genome-wide association studies; gene–environment interactions

INTRODUCTION

Depression is a common disorder with a complex etiology. There is a substantial heritable component to risk, although heritability estimates vary across severity and clinical subgroups [McGuffin et al., 1996; Sullivan et al., 2000; ten Doesschate et al., 2010; Lubke et al., 2012]. Recent genome-wide association studies have largely been unsuccessful in detecting genetic variants associated with depression with only one replicated finding [Ising et al., 2009; Sullivan et al., 2009; Lewis et al., 2010; Kohli et al., 2011; Shi et al., 2011; Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, 2012], in contrast to the relative success in other psychiatric and medical disorders [Ripke et al., 2011; Sklar et al., 2011]. One potential source of heterogeneity for genome-wide association studies is the strong role of environmental adversity in the etiology of depression.

Stressful life events (SLEs) such as illness, unemployment, or loss of a close relative or friend, have been strongly associated with the onset of depression [Paykel et al., 1980; Brown and Harris, 1989; Monroe et al., 1999]. Previously, SLEs have been exploited in the study of gene–environment interactions (G × E), which looked for genetic risk that varies according to the level of stress exposure [Caspi et al., 2003; Uher and McGuffin, 2008; Risch et al., 2009; Fisher et al., 2012]. However, G × E studies have low statistical power to capture the effect of genetic variants that would be associated with depression but whose effects may be overwhelmed at extreme levels of environmental exposure, or if the exposure to environmental factors is uneven in the population [Caspi et al., 2010]. Aside from statistical considerations, it is also relevant to put depressive episodes into the context of an individual's environment. Horwitz and Wakefield [2007] have argued that current classifications of major depression which, excluding bereavement, do not properly account for environmental factors in diagnosis. They suggest that ignoring the environmental context of depression risks medicalizing a rational response to stressful experiences. This may contribute to heterogeneity within cases by including those individuals reacting "normally" in response to their environment. On the other hand, restricting to those controls with exposure to environmental triggers for depression helps to screen out those genetically predisposed to the disorder that may develop it later in life when environment changes. For these reasons, G × E studies need to be complemented with a search for genetic effects that are associated with depression among individuals living in similar environments.

Psychiatric research is not the only area where matching or accounting for environmental measures in genome-wide association analyses is important, and so several methods have been proposed. An early example of this was the association between type 2 diabetes and *FTO*, a gene also associated with body mass index [Wellcome Trust Case Control Consortium, 2007]. The association with type 2 diabetes generally disappears in studies where only low body mass index cases were recruited, but was

apparent in those where cases were un-stratified [Frayling et al., 2007; Saxena et al., 2007; Freathy et al., 2008]. Several methodologies have been suggested for genome–environment wide interaction studies, or GEWIS, to combat the previously mentioned issues of low power in such studies [Thomas, 2010], but few focus on formal matching of samples based on environmental risk.

In order to disentangle environmental risk that might contaminate genetic studies, this article describes a novel method to incorporate environmental exposure variables into GWAS. Here we use propensity score matching to pair cases and controls on SLEs before performing a sign test across pairs to look for association of affection status with each single nucleotide polymorphism (SNP). As cases and controls differ in environmental exposure, this method reduces the number of controls with few environmental risk factors and decreases the number of cases with large environmental risk burdens. This allows the comparison of like with like between cases and controls that have experienced similar levels of environmental stress but developed contrasting outcomes, and potentially reduces heterogeneity within the sample.

METHODS

Sample

We used several studies that recruited participants in the same or similar manner and which are jointly known as RADIANT. The RADIANT study comprises three cohorts of depression cases: a case–control study, DeCC [Gaysina et al., 2008]; an affected sibling pair linkage study, DeNT [Farmer et al., 2004]; and a pharmacogenetic study, GENDEP [Uher et al., 2009]. DeCC cases were ascertained from three UK clinical centers (London, Cardiff, and Birmingham). Cases from DeNT and GENDEP were ascertained more widely from clinical centers across Europe for GENDEP (London, Brussels, Mannheim, Bonn, Brescia, Aarhus, Ljubljana, Poznan, and Zagreb) and from Europe and the US for DeNT (Birmingham, Cardiff, St. Louis, Aarhus, London, Bonn, Dublin, and Lausanne). Controls from DeCC and the bipolar case–control study, BACCS [Gaysina et al., 2009], were collected in the UK, and screened for absence of any psychiatric disorder. All DeCC and DeNT cases and 60.4% of GENDEP cases were recurrent (2 or more episodes of MDD). Genotyping was performed on the Illumina 610K beadarray. Individuals were excluded if their genotypic data showed a missing rate >1%, abnormal heterozygosity, a sex assignment that conflicted with phenotypic data, if they were related (up to second degree) with other study members, or were of non-European ancestry. SNPs with a low minor allele frequency (<1%) or showing departure from Hardy–Weinberg equilibrium ($P < 1 \times 10^{-5}$) were excluded. No imputation method was used. Full details of genotyping methods and quality control procedures for a genome-wide study of depression in the UK have been published [Lewis et al., 2010]. All study participants provided informed consent, and ethical approval was obtained from relevant institutional review boards. In all analyses we restricted to individuals recruited in Western European centers to reduce both potential genetic and reporting differences.

Stressful Life Event data

All studies included the List of Threatening Experiences Questionnaire [LTE-Q; Brugha et al., 1985; Brugha and Cragg, 1990] for both cases and controls, with cases and controls reporting for 6 months before interview and cases having additional information for 6 months before worst episode. For the purposes of this analysis only data for events prior to worst episode was used for cases. The LTE-Q was administered to all participants during their interview; this entailed asking whether they experienced an event, confirming the event occurred during the specified index period and obtaining some contextual information to establish that the reference event fulfilled the classification of the items listed on this instrument. For example, the respondent reporting a cold or flu would not be included under the "personal illness" category of event (insufficiently severe illness). While the original LTE-Q consists of 12 events, two of these were combined into a single item for the present study (these were: "did you have a separation due to marital difficulties?" and "did you break off a steady relationship?" which were combined into "did you have a separation due to marital difficulties or break off a steady relationship?"). Childbirth was added as an additional stressful life event. These events had previously been found to be associated with depression status, both as a cumulative score and individually [Hosang et al., 2010, 2012]. Each SLE was rated on a four-point scale severity, with a rating of not present, mild, moderate, or severe.

Matching

Cases and controls were first matched on ancestry-informative principal components (PCs) to deal with population stratification, and then independently matched on SLEs. Both steps were performed within STATA MP [StataCorp., 2011]. The first step used the CEM package [Iacus et al., 2009], which performs coarsened exact matching creating strata based on the desired variables. This program was developed to prune individuals that lacked closely matching pairs before using propensity scores to identify specific matches. Here it was used to create strata based on the first two PCs. Within each stratum, propensity score matching was used to match cases and controls based on each of the 12 items from the modified LTE-Q. This generated a probability of each individual being either a case or a control, based on these 12 SLEs. Affected individuals were then matched to an unaffected individual based on propensity scores using nearest neighbor matching in the psmatch2 package [Leuven and Sianesi, 2003]. A threshold difference in propensity score of 0.2 within each pair was imposed to restrict analysis to the best matched pairs, standard for this type of analysis [Guo and Fraser, 2009].

Association Testing: Logistic and Allelic Sign Test

A non-parametric test was performed for each SNP to assess whether cases carried more (or fewer) rare alleles than controls. For each SNP, the number of cases carrying more rare alleles than their paired control was tabulated, giving a total of n+ pairs. This was compared to the number of pairs where cases carried fewer rare

alleles than the paired controls (n−). Departure from the null hypothesis of an equal probability that the paired cases and controls carried more rare alleles was tested using a sign test (expected proportion of cases having more rare alleles than controls = (n+ + n−)/2).

For comparison, an unconditional logistic regression was performed within PLINK [Purcell et al., 2007] with analysis restricted to those cases and controls that were matched on both PCs and SLEs. In the logistic regression analysis, no PCs were used as covariates to assess the effect of matching on population stratification. To correct for multiple testing, a threshold *P*-value of 5E−8 was used to identify associations of genome-wide significance, with $P = 5E{-}6$ used as a cut off for suggestive significance.

RESULTS

Propensity Score Matching

Analysis of the first two principal components produced 86 strata. Pruning to only those strata with at least one case and control resulted in 3,497 individuals in 60 strata remaining from the initial 3,690. The number of individuals in each stratum ranged from 2 to 435, with a median number of 25 individuals. Within these individuals, the average score for stressful life events was 3.8 (SD 3.9) for cases and 1.3 (SD 1.8) for controls. Propensity score matching resulted in 805 pairs of cases and controls passing our requirements for thresholds on principal components and stressful life events. Within these individuals, both total number of SLEs and each individual SLE were no longer significantly associated with affected status ($P > 0.05$). The mean difference in propensity score was 0.022 (SE 0.0016). After matching, the average score for life events was 1.1 (SD 1.8) for cases and 1.1 (SD 1.7) for controls, showing that on average those cases with a very high number of environmental risk factors had not been matched and so removed. This matched cases with the highest genetic and lowest environmental component to controls (Table I).

Association Analysis

After quality control, 457,670 SNPs remained. Genome-wide analysis using a sign test of allele direction across all pairs resulted in a lambda value of 1.01, suggesting that matching within strata based on principal components effectively dealt with population stratification. There was a clear deflation in the number of SNPs with highly significant associations compared to the expected distribution of test statistics (Fig. 1), which may have been due to overcorrecting by matching on both SLEs and principal components. The sign test revealed no SNPs associated at suggestive or genome-wide significance (Fig. 2). Logistic regression was also performed in the restricted sample of matched individuals, without pairing. Again, no principal components were used and again after previous matching on PCs there was no evidence of population stratification (lambda = 1.02, Fig. 1). Two SNPs reached suggestive significance in this analysis (see Table II). The first was rs8050326 on the q24 arm of chromosome 16 near the gene *IRF8* ($P = 2.6 \times 10^{-7}$) which had been found in the original case–control analysis of this dataset. The other was rs11152166 on the q21 arm of chromosome 18 within the *CCBE1* gene ($P = 2.7 \times 10^{-6}$). Neither of these regions has been identified in depression studies outside of the primary analysis of this data. Correlation of the log *P* values from the sign test with those from the logistic regression was 0.25, suggesting they may have tested largely independent hypotheses with matching adding information to the analysis beyond simply pruning individuals.

TABLE I. Summary Statistics of Stressful Life Events by Affection Status

	Total sample		Matched sample	
Using % for events	**Cases**	**Controls**	**Cases**	**Controls**
N	2,626	1,064	805	805
Percentage female	71%	59%	71%	60%
Mean age[a] (SE)	37.8 (0.24)	41.5 (0.40)	38.29 (0.48)	41.9 (0.46)
Illness or injury of self	16.3	5.7	4.8	6.2
Illness or injury of other individual	17.5	15.0	12.3	12.9
Death of first degree relative	9.7	2.8	1.9	2.5
Death of other individual	12.3	15.3	9.9	10.2
Separation/divorce from partner	21.1	4.6	3.5	4.6
Conflict with close friend or family	24.1	7.6	6.0	7.9
Redundancy	8.5	1.9	1.0	1.9
Unemployment	8.6	3.6	1.7	2.6
Financial problems	13.8	3.8	2.9	3.5
Legal problems	3.0	1.0	0.4	0.6
Loss of valued item	5.9	5.4	2.5	3.3
Childbirth	7.0	0.8	0.4	0.7

[a]Age at interview for controls and age at worst episode for cases.

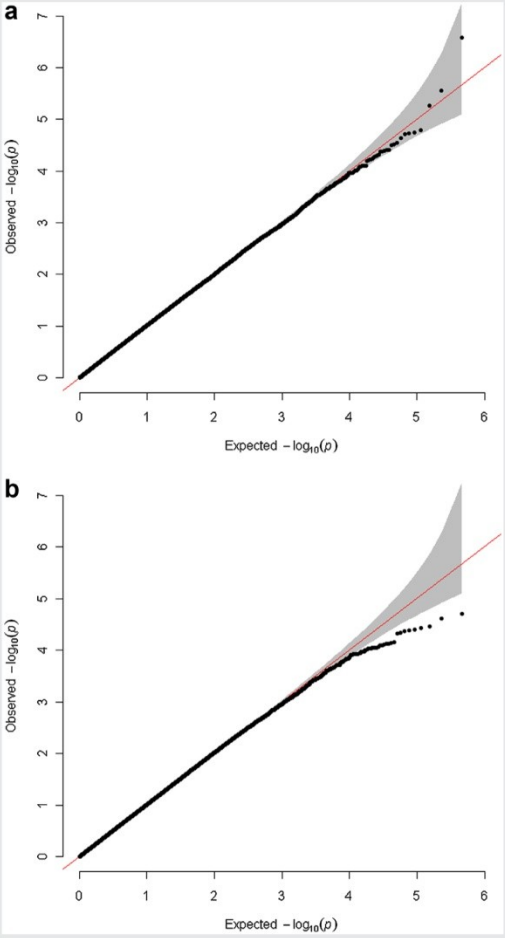


FIG 1. Comparison of QQ plots from genome-wide analysis of cases and controls matched on stressful life events for unpaired logistic regression (left, a) and paired sign test (right, b).

Figures 1 and 2 show a comparison of the QQ and Manhattan plots for these two analyses.

DISCUSSION

The role of environmental risk factors in depression is both important and complex. Differences between cases and controls in environmental risk could introduce a source of heterogeneity into genome-wide association studies for the disorder. Here we applied a novel approach to tackle this issue, restricting analysis to cases and controls matched for exposure to stressful life events. This is of particular importance as cases and controls differ in terms of exposure. While this is possibly partially due to reverse causation from traits like neuroticism that have been associated with both depression and increased reporting of negative events [Farmer et al., 2002], it also highlights the risk that differences in exposure to environmental risks may introduce heterogeneity into the sample. While our results did not provide any genome-wide significant results, the novel method for incorporating environmental data into GWAS may be more generally applicable.

The most likely explanation for the lack of positive findings here is a lack of power, considering the large sample sizes required to identify risk variants in genome-wide association studies. Propensity score matching on covariates that show a large difference between cases and controls, such as SLEs, will inevitably result in some individuals being unable to be matched. This reflects the large discrepancy between cases and controls on exposure to environmental adversity and illustrates the level of heterogeneity that may impair the power of GWAS of stress-related disorders. This method may be more applicable to studies where cases and controls have been ascertained to reduce differences in environment, or where the influence of covariates is believed to reduce power to a degree that is greater than that from the reduced sample size. This is plausible as it has been shown that power can be increased through conditioning cases and controls on the basis of clinical covariates [Zaitlen et al., 2012], while simply correcting for them can substantially reduce power unless the disorder is highly prevalent [Pirinen et al., 2012]. A further consideration in the merits of matching rather than covarying is how the covariate in question is distributed amongst cases and controls. If the distribution of a covariate does not overlap completely between cases and controls, as is true for higher burdens of SLEs, covarying out its effects becomes problematic.

An additional serious concern is that of an alternate framework for SLEs. It has been shown that the reporting of SLEs is itself heritable, including in the sample used for the present study [Plomin et al., 1990; Kendler and Baker, 2007; Power et al., 2012]. If this heritability overlaps considerably with the heritability of depression (plausible considering the association between depression and SLEs), then matching cases and controls on SLEs would increase their genetic similarity due to gene–environment correlation. Partial overlap between genetic factors contributing to depression and reported SLE has been reported, though only with respect to certain types of stressful life events [Thapar et al., 1998; Rice et al., 2003; Boardman et al., 2011]. Gene–environment correlation may have been problematic for another reason, as individuals were matched on both ancestry-informative principal components and SLEs. This may have led to over-correction, reducing the genetic differences between cases and controls. However, the presence of an underlying genetic component to SLEs only further strengthens the likelihood of genetically distinct sub-types of "environmental" and "biological" depression, and so the very heterogeneity this method was designed to remove. However, as only events in the previous 6 months were used, some environmental variance may have been left unaccounted for in our matching of cases and controls. Lastly, it is also possible our negative results are due to

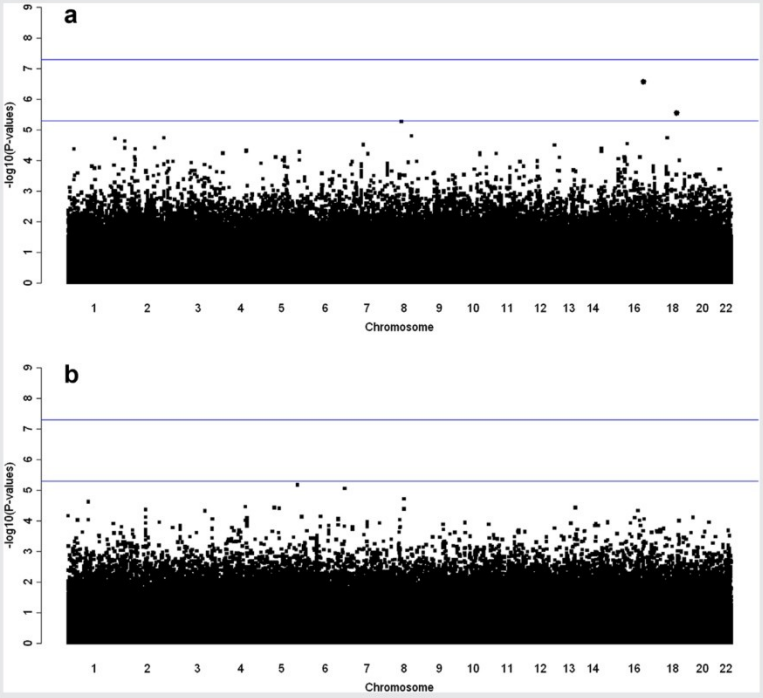


FIG 2. Comparison of Manhattan plots from genome-wide analysis of cases and controls matched on stressful life events for unpaired logistic regression (left, a) and paired sign test (right, b).

a small underlying genetic component of depression, though for a recurrent clinical cohort as examined here this should not be the case [McGuffin et al., 1996; Sullivan et al., 2000; Levinson, 2006]. As with any genome-wide association study it is possible that the causal variants are not tagged by those variants genotyped, for example, copy number variants or rare variants.

Our method of stratifying by principal components before matching on SLEs successfully removed evidence of population stratification in both the paired sign test and in the unpaired logistic regression analyses. However the main limitation of this study was clear, that the loss in power due to reduced sample size must be outweighed by the increased power from reduced heterogeneity. It's worth noting that our approach prunes sample size in proportion to the difference in environment between cases and controls, and so in proportion to the potential reduction of power this difference introduces. That there is such a large difference in SLEs between cases and controls identified here suggests the scale of potential heterogeneity introduced. We hope that by providing a detailed strategy for incorporating environmental factors into a GWAS will provide additional motivation to collect information on environmental adversity in future studies which may succeed in identifying genetic variants associated with depression and other stress-related disorder by accounting for environmental factors. This contrasts with recent large studies of psychiatric illness where the aim has

TABLE II. Top SNPs From Logistic Regression Analysis of Those Individuals Who Remained After Pruning for Principal Components and Stressful Life Events

CHR	SNP	BP	Effect allele	Frequency	OR	SE	L95	U95	*P*-value
16	rs8050326	8,46,61,643	A	0.44	1.45	0.072	1.26	1.67	2.64E−07
18	rs11152166	5,54,65,787	C	0.42	0.71	0.072	0.62	0.82	2.74E−06

been to maximize sample size, at the cost of standardized phenotype data [Purcell et al., 2007; Ripke et al., 2011; Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, 2012]. While this has succeeded with other psychiatric disorders, it has failed to yield genetic predictors for depression and so an alternate strategy of in-depth phenotyping should be pursued in tandem.

REFERENCES

Boardman JD, Alexander KB, Stallings MC. 2011. Stressful life events and depression among adolescent twin pairs. Biodemogr Soc Biol 57(1): 53–66.

Brown GW, Harris TO. 1989. Life events and illness. New York: Guildford Publications.

Brugha TS, Cragg D. 1990. The list of threatening experiences: The reliability and validity of a brief life events questionnaire. Acta Psychiatry Scand 82(1):77–81.

Brugha T, Bebbington P, Tennant C, Hurry J. 1985. The list of threatening experiences: A subset of 12 life event categories with considerable long-term contextual threat. Psychol Med 15(1):189–194.

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R. 2003. Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. Science 301(5631):386–389.

Caspi A, Hariri AR, Holmes A, Uher R, Moffitt TE. 2010. Genetic sensitivity to the environment: The case of the serotonin transporter gene and its implications for studying complex diseases and traits. Am J Psychiatry 167(5):509–527.

Consortium WTCC. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447(7145):661–678.

Consortium MDDWGotPG. 2013. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 18(4):497–511.

Farmer A, Redman K, Harris T, Mahmood A, Sadler S, Pickering A, McGuffin P. 2002. Neuroticism, extraversion, life events and depression. The Cardiff Depression Study. Br J Psychiatry 181:118–122.

Farmer A, Breen G, Brewster S, Craddock N, Gill M, Korszun A, Maier W, Middleton L, Mors O, Owen M, Perry J, Preisig M, Rietschel M, Reich T, Jones L, Jones I, McGuffin P. 2004. The Depression Network (DeNT) Study: Methodology and sociodemographic characteristics of the first 470 affected sibling pairs from a large multi-site linkage genetic study. BMC Psychiatry 4:42.

Fisher HL, Cohen-Woods S, Hosang GM, Uher R, Powell-Smith G, Keers R, Tropeano M, Korszun A, Jones L, Jones I, Owen M, Craddock N, Craig IW, Farmer AE, McGuffin P. 2012. Stressful life events and the serotonin transporter gene (5-HTT) in recurrent clinical depression. J Affect Disord 136(1–2):189–193.

Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch AM, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin MR, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJ, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CN, Doney AS, Morris AD, Smith GD, Hattersley AT, McCarthy MI. 2007. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316(5826):889–894.

Freathy RM, Timpson NJ, Lawlor DA, Pouta A, Ben-Shlomo Y, Ruokonen A, Ebrahim S, Shields B, Zeggini E, Weedon MN, Lindgren CM, Lango H, Melzer D, Ferrucci L, Paolisso G, Neville MJ, Karpe F, Palmer CN, Morris AD, Elliott P, Jarvelin MR, Smith GD, McCarthy MI, Hattersley AT, Frayling TM. 2008. Common variation in the FTO gene alters diabetes-related metabolic traits to the extent expected given its effect on BMI. Diabetes 57(5):1419–1426.

Gaysina D, Cohen S, Craddock N, Farmer A, Hoda F, Korszun A, Owen MJ, Craig IW, McGuffin P. 2008. No association with the 5,10-methylenetetrahydrofolate reductase gene and major depressive disorder: Results of the depression case control (DeCC) study and a meta-analysis. Am J Med Genet Part B 147B(6):699–706.

Gaysina D, Cohen-Woods S, Chow PC, Martucci L, Schosser A, Ball HA, Tozzi F, Perry J, Muglia P, Craig IW, McGuffin P, Farmer A. 2009. Association of the dystrobrevin binding protein 1 gene (DTNBP1) in a bipolar case–control study (BACCS). Am J Med Genet Part B 150B(6):836–844.

Guo SY, Fraser MW. 2009. Propensity score analysis: Statistical methods and applications (advanced quantitative techniques in the social sciences). Thousand Oaks, CA: Sage Publications.

Horwitz AV, Wakefield JC. 2007. The loss of sadness. Oxford, UK: Oxford University Press.

Hosang GM, Korszun A, Jones L, Jones I, Gray JM, Gunasinghe CM, McGuffin P, Farmer AE. 2010. Adverse life event reporting and worst illness episodes in unipolar and bipolar affective disorders: Measuring environmental risk for genetic research. Psychol Med 40(11):1829–1837.

Hosang GM, Korszun A, Jones L, Jones I, McGuffin P, Farmer AE. 2012. Life-event specificity: Bipolar disorder compared with unipolar depression. Br J Psychiatry 201(6):458–465.

Iacus SM, King G, Porro G. 2009. CEM: Coarsened exact matching software. J Stat Softw 30(9).

Ising M, Lucae S, Binder EB, Bettecken T, Uhr M, Ripke S, Kohli MA, Hennings JM, Horstmann S, Kloiber S, Menke A, Bondy B, Rupprecht R, Domschke K, Baune BT, Arolt V, Rush AJ, Holsboer F, Muller-Myhsok B. 2009. A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. Arch Gen Psychiatry 66(9):966–975.

Kendler KS, Baker JH. 2007. Genetic influences on measures of the environment: A systematic review. Psychol Med 37(5):615–626.

Kohli MA, Lucae S, Saemann PG, Schmidt MV, Demirkan A, Hek K, Czamara D, Alexander M, Salyakina D, Ripke S, Hoehn D, Specht M, Menke A, Hennings J, Heck A, Wolf C, Ising M, Schreiber S, Czisch M, Muller MB, Uhr M, Bettecken T, Becker A, Schramm J, Rietschel M, Maier W, Bradley B, Ressler KJ, Nothen MM, Cichon S, Craig IW, Breen G, Lewis CM, Hofman A, Tiemeier H, van Duijn CM, Holsboer F, Muller-Myhsok B, Binder EB. 2011. The neuronal transporter gene SLC6A15 confers risk to major depression. Neuron 70(2):252–265.

Leuven E, Sianesi B. 2003. PSMATC H2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Statistical Software Components: Boston College Department of Economics.

Levinson DF. 2006. The genetics of depression: A review. Biol Psychiatry 60(2):84–92.

Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, Weale ME, Schosser A, Paredes UM, Rivera M, Craddock N, Owen MJ, Jones L, Jones I, Korszun A, Aitchison KJ, Shi JX, Quinn JP, MacKenzie A, Vollenweider P, Waeber G, Heath S, Lathrop M, Muglia P, Barnes MR, Whittaker JC, Tozzi F, Holsboer F, Preisig M, Farmer AE, Breen G, Craig IW, McGuffin P. 2010. Genome-wide association study of major recurrent depression in the UK population. Am J Psychiatry 167(8):949–957.

Lubke GH, Hottenga JJ, Walters R, Laurin C, de Geus EJ, Willemsen G, Smit JH, Middeldorp CM, Penninx BW, Vink JM, Boomsma DI. 2012. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. Biol Psychiatry 72(8):707–709.

McGuffin P, Katz R, Watkins S, Rutherford J. 1996. A hospital-based twin register of the heritability of DSM-IV unipolar depression. Arch Gen Psychiatry 53(2):129–136.

Monroe SM, Rohde P, Seeley JR, Lewinsohn PM. 1999. Life events and depression in adolescence: Relationship loss as a prospective risk factor for first onset of major depressive disorder. J Abnorm Psychol 108 (4):606–614.

Paykel ES, Emms EM, Fletcher J, Rassaby ES. 1980. Life events and social support in puerperal depression. Br J Psychiatry 136:339–346.

Pirinen M, Donnelly P, Spencer CC. 2012. Including known covariates can reduce power to detect genetic effects in case–control studies. Nat Genet 44(8):848–851.

Plomin R, Lichtenstein P, Pedersen NL, McClearn GE, Nesselroade JR. 1990. Genetic influence on life events during the last half of the life span. Psychol Aging 5(1):25–30.

Power RA, Wingenbach T, Cohen-Woods S, Uher R, Ng MY, Butler AW, Ising M, Craddock N, Owen MJ, Korszun A, Jones L, Jones I, Gill M, Rice JP, Maier W, Zobel A, Mors O, Placentino A, Rietschel M, Lucae S, Holsboer F, Binder EB, Keers R, Tozzi F, Muglia P, Breen G, Craig IW, Muller-Myhsok B, Kennedy JL, Strauss J, Vincent JB, Lewis CM, Farmer AE, McGuffin P. 2012. Estimating the heritability of reporting stressful life events captured by common genetic variants. Psychol Med 1–7.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81(3):559–575.

Rice F, Harold GT, Thapar A. 2003. Negative life events as an account of age-related differences in the genetic aetiology of depression in childhood and adolescence. J Child Psychol Psychiatry 44(7):977–987.

Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA, Lin DY, Duan J, Ophoff RA, Andreassen OA, Scolnick E, Cichon S, Clair DS, Corvin A, Gurling H, Werge T, Rujescu D, Blackwood DHR, Pato CN, Malhotra AK, Purcell S, Dudbridge F, Neale BM, Rossin L, Visscher PM, Posthuma D, Ruderfer DM, Fanous A, Stefansson H, Steinberg S, Mowry BJ, Golimbet V, De Hert M, Jonsson EG, Bitter I, Pietilainen OPH, Collier DA, Tosato S, Agartz I, Albus M, Alexander M, Amdur RL, Amin F, Bass N, Bergen SE, Black DW, Borglum AD, Brown MA, Bruggeman R, Buccola NG, Byerley WF, Cahn W, Cantor RM, Carr VJ, Catts SV, Choudhury K, Cloninger CR, Cormican P, Craddock N, Danoy PA, Datta S, De Haan L, Demontis D, Dikeos D, Djurovic S, Donnelly P, Donohoe G, Duong L, Dwyer S, Fink-Jensen A, Freedman R, Freimer NB, Friedl M, Georgieva L, Giegling I, Gill M, Glenthoj B, Godard S, Hamshere M, Hansen M, Hansen T, Hartmann AM, Henskens FA, Hougaard DM, Hultman CM, Ingason A, Jablensky AV, Jakobsen KD, Jay M, Jurgens G, Kahn R, Keller MC, Kenis G, Kenny E, Kim Y, Kirov GK, Konnerth H, Konte B, Krabbendam L, Krasucki R, Lasseter VK, Laurent C, Lawrence J, Lencz T, Lerer FB, Liang KY, Lichtenstein P, Lieberman JA, Linszen DH, Lonnqvist J, Loughland CM, Maclean AW, Maher BS, Maier W, Mallet J, Malloy P, Mattheisen M, Mattingsdal M, McGhee KA, McGrath JJ, McIntosh A, McLean DE, McQuillin A, Melle I, Michie PT, Milanova V, Morris DW, Mors O, Mortensen PB, Moskvina V, Muglia P, Myin-Germeys I, Nertney DA, Nestadt G, Nielsen J, Nikolov I, Nordentoft M, Norton N, Nothen MM, O'Dushlaine CT, Olincy A, Olsen L, O'Neill FA, Orntoft TF, Owen MJ, Pantelis C, Papadimitriou G, Pato MT, Peltonen L, Petursson H, Pickard B, Pimm J, Pulver AE, Puri V, Quested D, Quinn EM, Rasmussen HB, Rethelyi JM, Ribble R, Rietschel M, Riley BP, Ruggeri M, Schall U, Schulze TG, Schwab SG, Scott RJ, Shi JX, Sigurdsson E, Silverman JM, Spencer CCA, Stefansson K, Strange A, Strengman E, Stroup TS, Suvisaari J, Terenius L, Thirumalai S, Thygesen JH, Timm S, Toncheva D, van den Oord E, van Os J, van Winkel R, Veldink J, Walsh D, Wang AG, Wiersma D, Wildenauer DB, Williams HJ, Williams NM, Wormley B, Zammit S, Sullivan PF, O'Donovan MC, Daly MJ, Gejman PV, Genome-Wide SP. 2011. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 43(10):969–976.

Risch N, Herrell R, Lehner T, Liang KY, Eaves L, Hoh J, Griem A, Kovacs M, Ott J, Merikangas KR. 2009. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression A meta-analysis. J Am Med Assoc 301(23):2462–2471.

Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S. 2007. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 316(5829):1331–1336.

Shi J, Potash JB, Knowles JA, Weissman MM, Coryell W, Scheftner WA, Lawson WB, DePaulo JR, Gejman PV, Sanders AR, Johnson JK, Adams P, Chaudhury S, Jancic D, Evgrafov O, Zvinyatskovskiy A, Ertman N, Gladis M, Neimanas K, Goodell M, Hale N, Ney N, Verma R, Mirel D, Holmans P, Levinson DF. 2011. Genome-wide association study of recurrent early-onset major depressive disorder. Mol Psychiatry 16(2):193–201.

Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N, Edenberg HJ, Nurnberger JI Jr, Rietschel M, Blackwood D, Corvin A, Flickinger M, Guan W, Mattingsdal M, McQuillin A, Kwan P, Wienker TF, Daly M, Dudbridge F, Holmans PA, Lin D, Burmeister M, Greenwood TA, Hamshere ML, Muglia P, Smith EN, Zandi PP, Nievergelt CM, McKinney R, Shilling PD, Schork NJ, Bloss CS, Foroud T, Koller DL, Gershon ES, Liu C, Badner JA, Scheftner WA, Lawson WB, Nwulia EA, Hipolito M, Coryell W, Rice J, Byerley W, McMahon FJ, Schulze TG, Berrettini W, Lohoff FW, Potash JB, Mahon PB, McInnis MG, Zollner S, Zhang P, Craig DW, Szelinger S, Barrett TB, Breuer R, Meier S, Strohmaier J, Witt SH, Tozzi F, Farmer A, McGuffin P, Strauss J, Xu W, Kennedy JL, Vincent JB, Matthews K, Day R, Ferreira MA, O'Dushlaine C, Perlis R, Raychaudhuri S, Ruderfer D, Hyoun PL, Smoller JW, Li J, Absher D, Thompson RC, Meng FG, Schatzberg AF, Bunney WE, Barchas JD, Jones EG, Watson SJ, Myers RM, Akil H, Boehnke M, Chambert K, Moran J, Scolnick E, Djurovic S, Melle I, Morken G, Gill M, Morris D, Quinn E, Muhleisen TW, Degenhardt FA, Mattheisen M, Schumacher J, Maier W, Steffens M, Propping P, Nothen MM, Anjorin A, Bass N, Gurling H, Kandaswamy R, Lawrence J, McGhee K, McIntosh A, McLean AW, Muir WJ, Pickard BS, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Jones IR, Kirov G, Moskvina V, Nikolov I, O'Donovan MC, Owen MJ, Collier DA, Elkin A, Williamson R, Young AH, Ferrier IN, Stefansson K, Stefansson H, Thornorgeirsson T, Steinberg S, Gustafsson O, Bergen SE, Nimgaonkar V, Hultman C, Landen M, Lichtenstein P, Sullivan P, Schalling M, Osby U, Backlund L, Frisen L, Langstrom N, Jamain S, Leboyer M, Etain B, Bellivier F, Petursson H, Sigur Sson E, Muller-Mysok B, Lucae S, Schwarz M, Schofield PR, Martin N, Montgomery GW, Lathrop M, Oskarsson H, Bauer M, Wright A, Mitchell PB, Hautzinger M, Reif A, Kelsoe JR, Purcell SM. 2011. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 43(10):977–983.

StataCorp. 2011. Stata statistical software: Release 12. College Station, TX: StataCorp. LP.

Sullivan PF, Neale MC, Kendler KS. 2000. Genetic epidemiology of major depression: Review and meta-analysis. Am J Psychiatry 157(10):1552–1562.

Sullivan PF, de Geus EJC, Willemsen G, James MR, Smit JH, Zandbelt T, Arolt V, Baune BT, Blackwood D, Cichon S, Coventry WL, Domschke K, Farmer A, Fava M, Gordon SD, He Q, Heath AC, Heutink P, Holsboer F, Hoogendijk WJ, Hottenga JJ, Hu Y, Kohli M, Lin D, Lucae S, MacIntyre DJ, Maier W, McGhee KA, McGuffin P, Montgomery GW, Muir WJ, Nolen WA, Nothen MM, Perlis RH, Pirlo K, Posthuma D, Rietschel M, Rizzu P, Schosser A, Smit AB, Smoller JW, Tzeng JY, van Dyck R, Verhage M, Zitman FG, Martin NG, Wray NR, Boomsma DI, Penninx BWJH. 2009. Genome-wide association for major depressive disorder: A possible role for the presynaptic protein piccolo. Mol Psychiatry 14(4):359–375.

ten Doesschate MC, Koeter MWJ, Bockting CLH, Schene AH, Group DS. 2010. Health related quality of life in recurrent depression: A comparison with a general population sample. J Affect Disord 120(1–3):126–132.

Thapar A, Harold G, McGuffin P. 1998. Life events and depressive symptoms in childhood-shared genes or shared adversity? A research note. J Child Psychol Psychiatry 39(8):1153–1158.

Thomas D. 2010. Gene–environment-wide association studies: Emerging approaches. Nat Rev Genet 11(4):259–272.

Uher R, McGuffin P. 2008. The moderation by the serotonin transporter gene of environmental adversity in the aetiology of mental illness: Review and methodological analysis. Mol Psychiatry 13(2):131–146.

Uher R, Maier W, Hauser J, Marusic A, Schmael C, Mors O, Henigsberg N, Souery D, Placentino A, Rietschel M, Zobel A, Dmitrzak-Weglarz M, Petrovic A, Jorgensen L, Kalember P, Giovannini C, Barreto M, Elkin A, Landau S, Farmer A, Aitchison KJ, McGuffin P. 2009. Differential efficacy of escitalopram and nortriptyline on dimensional measures of depression. Br J Psychiatry 194(3):252–259.

Zaitlen N, Lindstrom S, Pasaniuc B, Cornelis M, Genovese G, Pollack S, Barton A, Bickeboller H, Bowden DW, Eyre S, Freedman BI, Friedman DJ, Field JK, Groop L, Haugen A, Heinrich J, Henderson BE, Hicks PJ, Hocking LJ, Kolonel LN, Landi MT, Langefeld CD, Le Marchand L, Meister M, Morgan AW, Raji OY, Risch A, Rosenberger A, Scherf D, Steer S, Walshaw M, Waters KM, Wilson AG, Wordsworth P, Zienolddiny S, Tchetgen ET, Haiman C, Hunter DJ, Plenge RM, Worthington J, Christiani DC, Schaumberg DA, Chasman DI, Altshuler D, Voight B, Kraft P, Patterson N, Price AL. 2012. Informed conditioning on clinical covariates increases power in case–control association studies. PLoS Genet 8(11):e1003032.

Psychological Medicine (2013), **43**, 1965–1971. © Cambridge University Press 2012
doi:10.1017/S0033291712002589

ORIGINAL ARTICLE

Estimating the heritability of reporting stressful life events captured by common genetic variants

R. A. Power[1]*, T. Wingenbach[1], S. Cohen-Woods[1], R. Uher[1,2], M. Y. Ng[1], A. W. Butler[1,3], M. Ising[4], N. Craddock[5], M. J. Owen[5], A. Korszun[6], L. Jones[7], I. Jones[6], M. Gill[8], J. P. Rice[9], W. Maier[10], A. Zobel[10], O. Mors[11], A. Placentino[12], M. Rietschel[13], S. Lucae[4], F. Holsboer[4], E. B. Binder[4], R. Keers[1,6], F. Tozzi[14], P. Muglia[14,15,16], G. Breen[1,17], I. W. Craig[1], B. Müller-Myhsok[4], J. L. Kennedy[18,19], J. Strauss[18,19], J. B. Vincent[18,19], C. M. Lewis[1], A. E. Farmer[1] and P. McGuffin[1]

[1] *MRC Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, UK*; [2] *Department of Psychiatry, Dalhousie University, Halifax, NS, Canada*; [3] *Department of Psychiatry, University of Hong Kong, Hong Kong, Special Administrative Region, China*; [4] *Max Planck Institute of Psychiatry, Munich, Germany*; [5] *MRC Centre for Neuropsychiatric Genetics and Genomics, Neuroscience and Mental Health Research Institute, Cardiff University, UK*; [6] *Barts and The London Medical School, Queen Mary University of London, UK*; [7] *Department of Psychiatry, School of Clinical and Experimental Medicine, University of Birmingham, UK*; [8] *Department of Psychiatry, Trinity Centre for Health Science, Dublin, Ireland*; [9] *Department of Psychiatry, Washington University, St Louis, MO, USA*; [10] *Department of Psychiatry, University of Bonn, Germany*; [11] *Centre for Psychiatric Research, Aarhus University Hospital, Risskov, Denmark*; [12] *Psychiatric Unit 23, Department of Mental Health, Spedali Civili Hospital and Biological Psychiatry Unit, Centro San Giovanni di Dio, FBF, Brescia, Italy*; [13] *Division of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Mannheim, Germany*; [14] *GlaxoSmithKline Research and Development, Verona, Italy*; [15] *Department of Psychiatry, University of Toronto, Canada*; [16] *NeuroSearch A/S, Ballerup, Denmark*; [17] *NIHR Biomedical Research Centre for Mental Health, South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, King's College London, UK*; [18] *Neurogenetics Section, Centre for Addiction and Mental Health, Toronto, Canada*; [19] *Department of Psychiatry, University of Toronto, Canada*

Background. Although usually thought of as external environmental stressors, a significant heritable component has been reported for measures of stressful life events (SLEs) in twin studies.

Method. We examined the variance in SLEs captured by common genetic variants from a genome-wide association study (GWAS) of 2578 individuals. Genome-wide complex trait analysis (GCTA) was used to estimate the phenotypic variance tagged by single nucleotide polymorphisms (SNPs). We also performed a GWAS on the number of SLEs, and looked at correlations between siblings.

Results. A significant proportion of variance in SLEs was captured by SNPs (30%, $p=0.04$). When events were divided into those considered to be dependent or independent, an equal amount of variance was explained for both. This 'heritability' was in part confounded by personality measures of neuroticism and psychoticism. A GWAS for the total number of SLEs revealed one SNP that reached genome-wide significance ($p=4\times10^{-8}$), although this association was not replicated in separate samples. Using available sibling data for 744 individuals, we also found a significant positive correlation of $R^2=0.08$ in SLEs ($p=0.03$).

Conclusions. These results provide independent validation from molecular data for the heritability of reporting environmental measures, and show that this heritability is in part due to both common variants and the confounding effect of personality.

Received 15 July 2012; Revised 30 September 2012; Accepted 11 October 2012; First published online 14 December 2012

Key words: GCTA, heritability of environment, personality, stressful life events.

Introduction

As a better understanding of the complex interplay between genes and environment has emerged, some of the simpler notions of 'nature *versus* nurture' have fallen out of favour. Nowhere is this more obvious than in studies showing that many environmental measures in fact show substantial heritability (Plomin *et al*. 1990; Kendler & Baker, 2007). One such environmental factor is the adversity, or stressful life events (SLEs), an individual has encountered. SLEs have been known to play an important role in the development of mental disorders, for example as a trigger for the onset of major depression (Kendler *et al*. 1999*b*), psychosis (Bebbington *et al*. 1993; Kessing *et al*. 2004) and

* Address for correspondence: Mr R. A. Power, MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, London SE5 8AF, UK.
(Email: robert.r.power@kcl.ac.uk)

mania (Kennedy *et al*. 1983). Although SLEs might seem to be purely external environmental factors, there is evidence from twin studies for a genetic component ranging from 20% to 50% (Plomin *et al*. 1990; Kendler *et al*. 1993; Billig *et al*. 1996) when self-report measures are used. Longitudinal studies have also shown that patterns in reporting SLEs are consistent for individuals through life (Andrews, 1981; Woods *et al*. 1982). This heritability is often only found for SLEs that are considered 'dependent', that is potentially influenced by an individual's behaviour, suggesting that heritable aspects of an individual's behaviour may play a role in their choice of environment and exposure to SLEs, with 'independent' events often found to lack any heritable component (Kendler *et al*. 1999*a*; Bemmels *et al*. 2008; Boardman *et al*. 2011). Furthermore, there is some evidence that this heritability is, at least in part, mediated by factors such as personality (Saudino *et al*. 1997; Kandler *et al*. 2012). This may reflect gene–environment correlation, where genetically controlled behaviours make an individual more likely to encounter or seek out stressful environments (e.g. a propensity towards risk taking). Alternatively, this could reflect heritable personality types more likely to interpret or report events as stressful. This is supported by studies showing increased heritability of self-reported events compared to objective measures (Thapar & McGuffin, 1996; Kendler & Baker, 2007).

Although quantitative genetic investigations have shown heritability of SLEs, no molecular research has been carried out on known environmental measures to date. To expand upon the results of twin studies, we first aimed to look at the correlation between siblings in the reporting of SLEs to establish familiarity. Second, we estimated the additive genetic variance explained by single nucleotide polymorphisms (SNPs) for the reporting of SLEs, using the software tool genome-wide trait association analysis (GCTA; Yang *et al*. 2011). We also examined if these estimates of the heritability captured by common SNPs varied with regard to those events supposedly dependent or independent of an individual's actions, or after correcting for personality measures. Third, we sought to identify specific genetic variants associated with the reporting of SLEs in a genome-wide association study (GWAS).

Method

Samples

The sample consists of individuals from the Depression Case Control (DeCC) study (Cohen-Woods *et al*. 2009) and the Depression Network (DeNT) study (Farmer *et al*. 2004). Both were retrospective studies of DSM-IV/ICD-10-defined recurrent depression of at least moderate severity. The controls were recruited through the Medical Research Council (MRC) general practice research framework, and screened for any history of psychiatric disease. DeCC provided 2327 individuals (1222 cases and 1106 controls) from the UK with both phenotype and genotype data. DeNT provided 744 sibling pairs with phenotype data for correlation analysis, and 294 genotyped individuals for genetic analysis. Only UK individuals were used for genetic analyses to avoid population stratification. The studies were approved by the local ethical committees and informed written consent was obtained from all participants.

Measures

The SLE data were derived from the Brief Life Event Questionnaire (BLEQ), which consisted of 12 items and was based on the List of Threatening Experiences Questionnaire (LTE-Q; Brugha *et al*. 1985). Individuals had to state whether or not they experienced the respective SLEs within the past 6 months. As this was a retrospective study, cases were not selected for in episode. Response categories were defined in a binary fashion, either present or absent, and a sum of the number of reported SLEs was constructed for every individual. Following the LTE-Q categories, SLEs were split into those considered dependent on an individual's actions and those that were deemed to be independent. Illness, death and being robbed were considered as independent SLEs; unemployment, separation, financial problems and legal matters as SLEs dependent on individuals' behaviour. This totalled five independent events and seven dependent events. These sums of numbers of independent and dependent events were then used in secondary analyses.

Personality was evaluated using the revised Eysenck Personality Questionnaire (EPQ-R), with separate measures for extraversion, psychoticism and neuroticism (Eysenck & Eysenck, 1975; Eysenck *et al*. 1985). Each of these personality measures was used individually as a covariate to test for confounding in the heritability of reporting SLEs.

Genotyping and quality control

For the whole sample Illumina Human Hap610-Quad BeadChips [Centre National de Genotypage (CNG), France] were used for genotyping. The genotyping is described in more detail elsewhere (Lewis *et al*. 2010). Individuals were excluded if the missing rate of genotypic data was >1%, the data showed abnormal

heterozygosity or discrepant gender assignment, or the participants were close relatives (up to second degree) or of non-European ancestry. If SNPs showed minor allele frequency $<1\%$, missingness of $>1\%$ or departure from Hardy–Weinberg equilibrium ($p<1\times10^{-5}$), they were excluded.

Statistical analyses

Sibling correlation for SLEs

From the DeNT sample one sibling pair was selected from each family. Spearman's correlation was conducted in Stata (StataCorp, 2011) for all SLEs, and for independent and dependent SLEs separately. For this analysis we also created a sum of SLEs that were not defined by familial events; thus the items 'death of parent', 'serious illness, injury or assault of a close relative' and 'death of close family friend or another relative' were excluded.

GCTA

GCTA was implemented to investigate the phenotypic variance in the number of SLEs explained by all SNPs (Yang *et al*. 2011). Analyses were conducted for all SLEs and separately for those considered dependent or independent. Correction for case–control status, study, genotyping batch and the first five principal components was applied. The addition of further principal components did not seem to alter the findings. Individuals were also screened for relatedness, and in any pair where $R>0.025$, one individual was excluded. To test the effect of personality, we then repeated this analysis correcting for psychoticism, neuroticism and extraversion separately.

GWAS

To identify specific genetic variants related to the number of reported SLEs, GWAS were conducted for all SLEs, and for dependent and independent SLEs separately, using PLINK (Purcell *et al*. 2007). The threshold for genome-wide significance was set to $p<5\times10^{-8}$, and $p<5\times10^{-6}$ was used for suggestive significance (Dudbridge & Gusnanto, 2008). The first two ancestry informative principal components were included as covariates and no inflation of genomic control values above $\lambda=1.05$ was seen.

Replication

Top hits from the GWAS were replicated within two separate samples. The first was a cohort of controls from a bipolar dataset from Toronto, collected as part of the bipolar disorder arm of the RADIANT studies, to which DeCC and DeNT belong (Gaysina *et al*. 2009). They totalled 257 individuals with phenotype and genotype information, collected in a similar manner to the DeCC and DeNT studies. The second was a set of 894 cases from a clinical depression cohort of individuals in Munich, again collected in a similar manner to the DeCC and DeNT studies (Tozzi *et al*. 2008; Muglia *et al*. 2010). For the Munich sample, individual event data were not available, only a summary score for the BLEQ for cases. Therefore, this sample could be used for replicating findings for the total number of events, but not hits specific to dependent or independent events.

Results

Sibling correlation for SLEs

Altogether, 744 sibling pairs were available to test the familiality of SLEs. Spearman's correlation was conducted to investigate the relationship between the number of SLEs reported by siblings. Significant positive correlations were found for all variables, with the total number of events correlating at 0.19 ($p<0.001$). Both independent and dependent events were correlated, although independent events to a much greater degree (an R^2 of 0.25 compared to 0.08, with $p<0.001$ and $p=0.03$ respectively). This is probably because independent events included many that were related to illness or death of family members and an analysis of total events, excluding those specifically family related, produced results similar to just dependent events ($R^2=0.08$, $p=0.03$).

GCTA

After quality control, genome-wide association data were available from 2578 unrelated individuals and 541628 SNPs. The composition of these samples is outlined in Table 1, with the average number of events in the past 6 months prior to interview per person near to one in all. The correlation between total number of independent and dependent events was 0.15. The results from the GCTA show that a significant proportion of variance could be attributed to common genetic variants for all, dependent and independent SLEs. The phenotypic variance accounted for by all SNPs for the number of reported SLEs was 29% ($p=0.03$, S.E. $=0.16$). When distinguishing between SLEs dependent on and independent of subjects' behaviour, SNPs explained 30% of the variance ($p=0.03$, S.E. $=0.16$) for dependent events and 26% ($p=0.04$, S.E. $=0.15$) for independent events. The GCTA was rerun after correcting for the effects of personality, giving the results outlined in Table 2. The total variance explained by SNPs for the number of SLEs was

Table 1. *Description of studies and the distribution of stressful life events (SLEs)*

Study	Genotyped UK individuals	Cases (%)	Female (%)	Mean age (years)	Mean number of SLEs (S.E.)
DeCC	2327	52.5	64.7	44.3	1.0 (0.26)
DeNT	294	100	76.2	45.6	1.3 (0.09)
Toronto	257	0	52.1	41.8	1.1 (0.08)
Munich	864	100	67.3	50.8	2.3 (0.11)[a]

DeCC, Depression Case Control study; DeNT, Depression Network study; S.E., standard error.

[a] Note that for the Munich replication sample, the mean Brief Life Event Questionnaire (BLEQ) score is displayed rather than the number of events.

no longer found to be significant after correcting for either neuroticism (from 29% to 23%, $p=0.07$) or psychoticism (to 17%, $p=0.15$). Neuroticism scores did not affect GCTA estimates when events were separated into those considered dependent on or independent of an individual's actions, and psychoticism only confounded the analysis of dependent events. Accounting for extraversion scores did not affect the results.

GWAS

For the analysis of the number of SLEs as a continuous trait in 2578 cases, the study had 50% power to detect association with an SNP with a frequency of 25%, accounting for 1.1% variation in SLEs, or 80% power for 1.5% variation. No genome-wide significant results were found for the total number of events in the genome-wide association, although five SNPs reached suggestive significance. One of these (rs4927134 on chromosome 1) reached genome-wide significance when only dependent events were included. One SNP not found in the primary analysis of all events (rs16837293) reached suggestive significance when only dependent events were included, and one of the SNPs in the primary analysis (rs17040523) was found to be in suggestive significance when using only independent events. Details of the SNPs with the strongest association found are given in Table 3. None of these top hits reached even nominal significance in the replication samples ($p<0.05$).

Discussion

This study has provided new molecular evidence showing that common genetic variants explain a significant proportion of the variance in self-reported environmental factors. Twin studies have previously suggested a heritability of approximately 20–50% for the number of reported SLEs (Plomin *et al*. 1990; Kendler *et al*. 1993; Billig *et al*. 1996) and, in keeping with this, we found that ~30% of the phenotypic variance in SLEs was tagged by common SNPs. A degree of underestimation is expected in this study as only common single nucleotide variants were genotyped, often relying on linkage disequilibrium to tag the true number of common variants, and missing copy number or rare variants entirely. This may suggest that the 'true' heritability is towards the top end of the range found in twin studies. Of note, there was no clear difference in the variance explained for supposedly independent or dependent SLEs, although the confidence intervals are large. This is in contrast to the findings from twin studies, but may reflect the fairly crude division of independent and dependent events in this analysis.

Regarding the role of personality, correcting for either neuroticism or psychoticism resulted in a non-significant heritability of the total number of SLEs, with psychoticism also showing a confounding effect when analysis was restricted to dependent events. Psychoticism has previously been associated with increased reckless and impulsive behaviour (Pickering *et al*. 2003), which may impact the risk of those events likely to be labelled 'dependent' to a greater extent. Neuroticism has been associated with a high level of depression symptoms (Farmer *et al*. 2002), and perhaps its effect captured a difference in the sample due to inclusion of depressed individuals that was not fully corrected for. This highlights one of the main limitations of this study, which is the inclusion of cases and controls. Despite correcting for affected status and using the number of events at 6 months before interview rather than before episode, the results may not be entirely reflective of a sample from the general population. Unfortunately, because of the need for large samples, the heritability in controls alone could not be tested. These results are, however, in line with findings from twin and family studies, showing that neuroticism accounts for a portion of the heritability

Table 2. *Variance explained by common SNPs for total stressful life events (SLEs), and split into those considered dependent or independent of an individual's state of mind. Also shown are the adjusted estimates after correcting for the personality measures neuroticism, extraversion and psychoticism separately*

				Neuroticism		Extraversion		Psychoticism	
Events	Heritability	S.E.	*p* value	h^2	*p* value	h^2	*p* value	h^2	*p* value
All events	0.29	0.16	0.03	0.23	0.07	0.28	0.04	0.17	0.15
Dependent	0.30	0.16	0.03	0.29	0.04	0.31	0.03	0.20	0.11
Independent	0.26	0.15	0.04	0.27	0.04	0.28	0.04	0.27	0.04

SNP, Single nucleotide polymorphism; SLE, stressful life event; S.E., standard error.

Table 3. *SNPs with the strongest signals for associations in GWAS of the number of reported SLEs, and their associations in the two replication samples: controls from a bipolar sample in Toronto and cases from a depression study in Munich (which lacked information on whether the SLEs were independent or dependent)*

					RADIANT ($n=2578$)		Toronto ($n=257$)		Munich ($n=894$)		
Events	CHR	SNP	BP	Risk allele	β	*P*	β	*p*	β	*p*	Sign
All	1	rs4927134	54 756 695	C	0.28	5.8×10^{-8}	0.20	0.25	0.13	0.58	+ + +
All	7	rs758938	31 765 832	A	−0.16	1.2×10^{-6}	−0.12	0.31	0.14	0.33	− − +
All	2	rs17040523	2 874 200	T	0.26	1.3×10^{-6}	0.22	0.22	0.05	0.83	+ + +
All	7	rs11980485	31 779 361	G	−0.16	1.4×10^{-6}	−0.12	0.30	0.13	0.37	− − +
All	1	rs10494809	198 398 707	T	0.18	3.0×10^{-6}	−0.14	0.27	0.35	0.05	+ − +
Dependent	**1**	**rs4927134**	**54 756 695**	**C**	**0.20**	$\mathbf{4.1\times10^{-8}}$	**0.09**	**0.50**	–	–	+ +
Dependent	3	rs16837293	127 418 611	G	0.27	2.9×10^{-6}	−0.28	0.19	–	–	+ −
Independent	2	rs17040523	2 874 200	T	0.16	5.5×10^{-7}	0.11	0.24	–	–	+ +

SNP, Single nucleotide polymorphism; GWAS, genome-wide association study; SLE, stressful life event; CHR, chromosome; BP, base position; Sign, direction of effect per respective study.

of reporting stressful events (Dudbridge & Gusnanto, 2008; Kandler *et al*. 2012).

The association analysis of specific SNPs provided no replicable findings, but this is not entirely surprising because of the small effect sizes reported in associations with even the clearest of phenotypes. None of the SNPs found at suggestive significance were associated with personality disorders or other candidate genes. A strong positive correlation for the number of reported SLEs among siblings was also found, with the highest correlation for independent SLEs. This finding is self-explanatory as life events were included that overlap because of relatedness, for example 'serious illness of close relative', although excluding such items still leads to a significant positive correlation. This could be evidence for the role of genetic similarity, but of course could be due to shared environmental factors.

Taking together the results from the GWAS, GCTA and sibling comparison, we add to the now considerable evidence that the reporting of SLEs is heritable. Our results also support studies showing that at least part of this heritability can be attributed to confounding from heritable personality measures. It should be noted, however, that in this study we relied entirely on self-reporting, and using such 'soft' measures meant we were unable to differentiate between the heritability of propensity towards reporting SLEs and the heritability of experiencing SLEs. Nonetheless, this finding is of relevance to studies exploring gene–environment interactions, where gene–environment correlation may be a concern. It also provides more insight into the complexities of disentangling the heritable component of a disorder, such as depression, that is often associated with stressful events.

Acknowledgements

The RADIANT studies were supported by a joint grant from the MRC and GlaxoSmithKline (G0701420), and received financial support from the National Institute for Health Research (NIHR) Biomedical

Research Centre for Mental Health at the South London and Maudsley National Health Service (NHS) Foundation Trust and the Institute of Psychiatry, King's College London. The MARS project is funded by the German Federal Ministry of Education and Research (BMBF, project nos 01ES0811 and 01KG0709), and genome-wide genotyping was supported by the Bavarian Ministry of Commerce and the Excellence Foundation for the Advancement of the Max Planck Society. This work was also supported by the BMBF within the context of the German National Genome Research Network (NGFN-2 and NGFN-plus).

Declaration of Interest

K. J. Aitchison, A. E. Farmer and P. McGuffin have received consultancy fees and honoraria for participating in expert panels for pharmaceutical companies including GlaxoSmithKline. K. J. Aitchison declares interests through Advisory Boards for Johnson & Johnson, Lundbeck, Roche Diagnostics and Bristol–MyersSquibb; membership of the Bristol–Myers–Squibb UK Steering Group, 2003 to present; consultancy work for Roche Diagnostics, Johnson & Johnson Pharmaceutical Research and Development, Lundbeck, and Bristol–MyersSquibb Pharmaceuticals Limited; grants awarded by Johnson & Johnson Pharmaceutical Research and Development, Bristol-Myers Squibb Pharmaceuticals Limited, and E. Merck Pharmaceuticals. F. Tozzi and P. Muglia were employees of GlaxoSmithKline when the research was performed. M. Ising has received consultancy honoraria from MSD Merck. E. B. Binder has received grant support from PharmaNeuroboost.

References

Andrews G (1981). A prospective study of life events and psychological symptoms. *Psychological Medicine* **11**, 795–801.

Bebbington P, Wilkins S, Jones P, Foerster A, Murray R, Toone B, Lewis S (1993). Life events and psychosis. Initial results from the Camberwell Collaborative Psychosis Study. *British Journal of Psychiatry* **162**, 72–79.

Bemmels HR, Burt SA, Legrand LN, Iacono WG, McGue M (2008). The heritability of life events: an adolescent twin and adoption study. *Twin Research and Human Genetics* **11**, 257–265.

Billig JP, Hershberger SL, Iacono WG, McGue M (1996). Life events and personality in late adolescence: genetic and environmental relations. *Behavior Genetics* **26**, 543–554.

Boardman JD, Alexander KB, Stallings MC (2011). Stressful life events and depression among adolescent twin pairs. *Biodemography and Social Biology* **57**, 53–66.

Brugha T, Bebbington P, Tennant C, Hurry J (1985). The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. *Psychological Medicine* **15**, 189–194.

Cohen-Woods S, Gaysina D, Craddock N, Farmer A, Gray J, Gunasinghe C, Hoda F, Jones L, Knight J, Korszun A, Owen MJ, Sterne A, Craig IW, McGuffin P (2009). Depression Case Control (DeCC) Study fails to support involvement of the muscarinic acetylcholine receptor M2 (CHRM2) gene in recurrent major depressive disorder. *Human Molecular Genetics* **18**, 1504–1509.

Dudbridge F, Gusnanto A (2008). Estimation of significance thresholds for genomewide association scans. *Genetic Epidemiology* **32**, 227–234.

Eysenck HJ, Eysenck SBG (1975). *Manual of the Eysenck Personality Questionnaire*. Hodder and Stoughton: Sevenoaks, Kent.

Eysenck SBG, Eysenck HJ, Barrett P (1985). A revised version of the psychoticism scale. *Personality and Individual Differences* **6**, 21–29.

Farmer A, Breen G, Brewster S, Craddock N, Gill M, Korszun A, Maier W, Middleton L, Mors O, Owen M, Perry J, Preisig M, Rietschel M, Reich T, Jones L, Jones I, McGuffin P (2004). The Depression Network (DeNT) Study: methodology and sociodemographic characteristics of the first 470 affected sibling pairs from a large multi-site linkage genetic study. *BMC Psychiatry* **4**, 42.

Farmer A, Redman K, Harris T, Mahmood A, Sadler S, Pickering A, McGuffin P (2002). Neuroticism, extraversion, life events and depression. The Cardiff Depression Study. *British Journal of Psychiatry* **181**, 118–122.

Gaysina D, Cohen-Woods S, Chow PC, Martucci L, Schosser A, Ball HA, Tozzi F, Perry J, Muglia P, Craig IW, McGuffin P, Farmer A (2009). Association of the dystrobrevin binding protein 1 gene (DTNBP1) in a bipolar case-control study (BACCS). *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics* **150B**, 836–844.

Kandler C, Bleidorn W, Riemann R, Angleitner A, Spinath FM (2012). Life events as environmental states and genetic traits and the role of personality: a longitudinal twin study. *Behavior Genetics* **42**, 57–72.

Kendler KS, Baker JH (2007). Genetic influences on measures of the environment: a systematic review. *Psychological Medicine* **37**, 615–626.

Kendler KS, Karkowski LM, Prescott CA (1999*a*). The assessment of dependence in the study of stressful life events: validation using a twin design. *Psychological Medicine* **29**, 1455–1460.

Kendler KS, Karkowski LM, Prescott CA (1999*b*). Causal relationship between stressful life events and the onset of major depression. *American Journal of Psychiatry* **156**, 837–841.

Kendler KS, Neale M, Kessler R, Heath A, Eaves L (1993). A twin study of recent life events and difficulties. *Archives of General Psychiatry* **50**, 789–796.

Kennedy S, Thompson R, Stancer HC, Roy A, Persad E (1983). Life events precipitating mania. *British Journal of Psychiatry* **142**, 398–403.

Kessing LV, Agerbo E, Mortensen PB (2004). Major stressful life events and other risk factors for first admission with mania. *Bipolar Disorders* **6**, 122–129.

Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, Weale ME, Schosser A, Paredes UM, Rivera M, Craddock N, Owen MJ, Jones L, Jones I, Korszun A, Aitchison KJ, Shi JX, Quinn JP, MacKenzie A, Vollenweider P, Waeber G, Heath S, Lathrop M, Muglia P, Barnes MR, Whittaker JC, Tozzi F, Holsboer F, Preisig M, Farmer AE, Breen G, Craig IW, McGuffin P (2010). Genome-wide association study of major recurrent depression in the UK population. *American Journal of Psychiatry* **167**, 949–957.

Muglia P, Tozzi F, Galwey NW, Francks C, Upmanyu R, Kong XQ, Antoniades A, Domenici E, Perry J, Rothen S, Vandeleur CL, Mooser V, Waeber G, Vollenweider P, Preisig M, Lucae S, Muller-Myhsok B, Holsboer F, Middleton LT, Roses AD (2010). Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. *Molecular Psychiatry* **15**, 589–601.

Pickering A, Farmer A, Harris T, Redman K, Mahmood A, Sadler S, McGuffin P (2003). A sib-pair study of psychoticism, life events and depression. *Personality and Individual Differences* **34**, 613–623.

Plomin R, Lichtenstein P, Pedersen NL, McClearn GE, Nesselroade JR (1990). Genetic influence on life events during the last half of the life span. *Psychology and Aging* **5**, 25–30.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–575.

Saudino KJ, Pedersen NL, Lichtenstein P, McClearn GE, Plomin R (1997). Can personality explain genetic influences on life events? *Journal of Personality and Social Psychology* **72**, 196–206.

StataCorp (2011). *Stata Statistical Software: Release 12*. StataCorp LP: College Station, TX.

Thapar A, McGuffin P (1996). Genetic influences on life events in childhood. *Psychological Medicine* **26**, 813–820.

Tozzi F, Prokopenko I, Perry JD, Kennedy JL, McCarthy AD, Holsboer F, Berrettini W, Middleton LT, Chilcoat HD, Muglia P (2008). Family history of depression is associated with younger age of onset in patients with recurrent depression. *Psychological Medicine* **38**, 641–649.

Woods NF, Dery GK, Most A (1982). Stressful life events and perimenstrual symptoms. *Journal of Human Stress* **8**, 23–31.

Yang JA, Lee SH, Goddard ME, Visscher PM (2011). GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics* **88**, 76–82.

Molecular Psychiatry (2014), 1–4

www.nature.com/mp

ORIGINAL ARTICLE

Genetic predisposition to schizophrenia associated with increased use of cannabis

RA Power[1,2], KJH Verweij[3], M Zuhair[1], GW Montgomery[4], AK Henders[4], AC Heath[5], PAF Madden[5], SE Medland[4], NR Wray[2] and NG Martin[4]

Cannabis is the most commonly used illicit drug worldwide. With debate surrounding the legalization and control of use, investigating its health risks has become a pressing area of research. One established association is that between cannabis use and schizophrenia, a debilitating psychiatric disorder affecting ~1% of the population over their lifetime. Although considerable evidence implicates cannabis use as a component cause of schizophrenia, it remains unclear whether this is entirely due to cannabis directly raising risk of psychosis, or whether the same genes that increases psychosis risk may also increase risk of cannabis use. In a sample of 2082 healthy individuals, we show an association between an individual's burden of schizophrenia risk alleles and use of cannabis. This was significant both for comparing those who have ever versus never used cannabis ($P = 2.6 \times 10^{-4}$), and for quantity of use within users ($P = 3.0 \times 10^{-3}$). Although directly predicting only a small amount of the variance in cannabis use, these findings suggest that part of the association between schizophrenia and cannabis is due to a shared genetic aetiology. This form of gene–environment correlation is an important consideration when calculating the impact of environmental risk factors, including cannabis use.

Molecular Psychiatry advance online publication, 24 June 2014; doi:10.1038/mp.2014.51

INTRODUCTION

During the last quarter of the 20th century, cannabis use has increased to become the most widely used illicit drug in the world.[1] It is well established that cannabis use is much higher among schizophrenic patients than in the general population.[2] Cannabis intoxication can lead to an acute transient psychotic episode and produce short-term exacerbations of pre-existing psychotic symptoms,[3–5] an association that has been confirmed through the experimental administration of tetrahydrocannabinol.[6,7] Meta-analyses of prospective studies have found that cannabis use increases the likelihood of developing a psychotic illness by a factor of roughly two.[8–11] A dose response effect has been demonstrated,[12–14] and use in adolescence has been associated with the greatest risk.[15] Given the large health burden from schizophrenia and other psychotic disorders,[16] the view that cannabis use is a component cause of schizophrenia has heavily influenced discussion over the legislation surrounding cannabis use.

However, the relationship between schizophrenia and cannabis use may be more complicated than it initially seems. Despite a clear association between the two, the possibility of reverse causation has not been entirely excluded. Some small studies have suggested that it is in fact psychosis that is a risk factor for cannabis use, as those on a psychotic spectrum are more likely to experiment with drugs.[17,18] The strongest evidence comes from Ferdinand *et al.*[19] who found that the association was bidirectional, as cannabis-naive children with prodromal psychotic episodes had greater incidence of later cannabis use. However, a similarly sized study failed to replicate this finding.[20] There is also the possibility of attempts by patients at self-medication, as it has been suggested that cannabis use can reduce negative and affective symptoms in patients with an established psychotic disorder.[21–23]

The issue is further complicated by tentative evidence for interactions between cannabis use and genetic risk variants for schizophrenia.[24] Schizophrenia is known to be highly heritable with up to 80% of the variance explained by additive genetic effects,[25] and as sample sizes have increased a growing number of genetic risk variants have been identified.[26,27] Interactions between risk variants and cannabis use might explain why some individuals experience psychosis while others do not. However, cannabis use itself has been reported to be heritable,[28–30] although no genetic risk variants have been identified.[31] It is unclear to what extent the heritability of cannabis use results from shared heritability with other behavioural phenotypes such as schizophrenia predicting its use.

Here we test for such genetic overlap directly, and aim to discern the direction of causation between cannabis use and schizophrenia. Within a sample of 2082 healthy individuals, we tested to see whether polygenic risk scores for schizophrenia predict cannabis use. Polygenic risk scores reflect the cumulative burden of risk alleles carried by an individual as identified in a previous genome-wide association study (GWAS),[32] here of 13 833 schizophrenia cases and 18 310 controls.[27] Such an association with cannabis use would suggest that those genetically predisposed to schizophrenia use cannabis more frequently. This would mean that the association between schizophrenia and cannabis use is not simply one of an environmental risk factor, but rather involves gene–environment correlation, as individuals

[1]MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, UK; [2]Queensland Brain Institute, The University of Queensland, St Lucia, QLD, Australia; [3]Department of Developmental Psychology and EMGO Institute for Health and Care Research, VU University, Amsterdam, The Netherlands; [4]QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia and [5]Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA. Correspondence: R Power, MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, DeCrespigny Park, Denmark Hill, London SE5 8AF, UK.
E-mail: robert.r.power@kcl.ac.uk
Received 22 January 2014; revised 18 March 2014; accepted 22 April 2014

choose and shape their own environment based on their own innate preferences.

MATERIALS AND METHODS

The data used in this study come from the Australian Twin Registry. Data were obtained from two studies in which twins and their families participated in semi-structured diagnostic telephone interviews aimed primarily at assessing psychiatric health. Informed consent was obtained from all participants.

Sample 1 consisted of 6265 individuals aged between 23 and 39 years (mean = 29.9 ± 2.5) interviewed between 1996 and 2000. Participants were members of the young adult cohort, a volunteer panel of twins born between 1964 and 1971. The interview was based on a modified version of the SSAGA (Semi-Structured Assessment of the Genetics of Alcoholism[33]). Detailed information about the sample recruitment, the study procedure and the measures can be found elsewhere.[34] Sample 2 comprised 9688 individuals aged between 18 and 91 years (mean = 46.3 ± 11.3) interviewed between 2001 and 2005. Participants were members of the older and younger adult cohort of Australian twin pairs (born between 1895 and 1964, and between 1964 and 1971, respectively). A subset of this sample was ascertained based on large sibship size, or having a relative with nicotine or alcohol dependence. The interview used for this sample was also based on a modified version of the SSAGA. Further details about the sample and assessment can be found in Heath *et al*.[35]

A subset of the participants ($N = 1866$; 11.7%) participated in both studies, in which case we used data from the last assessment. The combined phenotypic sample consisted of 14 087 individuals, of whom 7172 were genotyped. In both studies, twins were asked the same items about cannabis use: (1) did you ever use marijuana?, (2) how old were you the very first time you tried marijuana (not counting the times you took it as prescribed)? and (3) how many times in your life have you used marijuana (do not count times when you used a drug prescribed for you and took the prescribed dose). Ever use was measured on a dichotomous scale (ever versus never), whereas age at initiation and quantity of use were open questions. Table 1 shows the prevalence of cannabis use for individuals included in the present study.

Genotype data were obtained using three different Illumina single nucleotide polygmorphism (SNP) genotyping platforms (317K, HumanCNV370-Quadv3, Human CNV370v1 and Human610-Quad). Standard quality control procedures were applied as outlined previously,[36] including checks for ancestry outliers, Hardy–Weinberg equilibrium ($P < 10^{-6}$), Mendelian errors, call rate, genotypic missingness (>5%), individual missingness (>5%) and minor allele frequency (< 0.01). Individuals were pruned on relatedness, removing one individual from each pair with relatedness >0.05, as determined from genetic data. The final sample therefore comprised 2082 'unrelated' individuals (see Table 1 for sample details).

Polygenic risk scores were constructed using the *P*-values and $\log_{10}$ odds ratios from the most recent large GWAS of schizophrenia, a meta-analysis of the Psychiatric Genomics Consortium's studies with additional Swedish samples totalling 13 833 cases and 18 310 controls.[27] SNPs were pruned for linkage disequilibrium using *P*-value informed clumping in PLINK,[37] using a cutoff of $R^2 = 0.25$ within 200 kb window. The major histocompatibility complex region of the genome was excluded, due to its complex linkage disequilibrium structure. After linkage disequilibrium pruning, 147 830 SNPs remained. Multiple scores were generated for each individual using the PLINK score option and based on top SNPs from the schizophrenia GWAS using varying significance thresholds ($P = 0.0001$, 0.001, 0.01. 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 and 1.0). Polygenic risk scores were tested for association with a binary ever versus never used cannabis and two quantitative traits for quantity of use and age at first use, in logistic and linear regressions, respectively. These analyses were corrected for the first 10 ancestry-informative principal components, genotyping platform, sex, age, age squared and sex by age. Analysis was performed in STATA.[38]

Table 1. Summary statistics of sample for cannabis use traits

	Users	*Non-users*
N	1011	1071
Mean age (s.e.)	41.3 (0.23)	53.0 (0.37)
Percentage female (%)	46.5	56.0
Mean age at initiation (s.e.)	19.6 (0.06)	—
Mean number of uses over lifetime (s.e.)	62.7 (4.56)	—

RESULTS

After pruning, 2082 unrelated individuals remained in our sample with both genotype and phenotype measures. Within the sample, 1011 individuals (48.6%) had ever used cannabis, of whom 997 had data on quantity of use. Mean number of usages of cannabis over lifetime was 62.7 (95% CI 53.8–71.6), and the mean age of initiation of use was 20.1 (95% CI 19.7–20.5). Males showed higher rates of use than females, 53.5% compared with 43.9% ($P < 0.001$), although no significant difference in age at initiation. Table 1 shows the summary statistics for the sample.

Polygenic risk scores for schizophrenia showed positive associations for ever versus never use of cannabis across all *P*-value thresholds, with the strongest association for those SNPs with *P*-values of 0.01 or below in the original schizophrenia GWAS (see Figure 1, $R^2 = 0.47\%$, $P = 2.6 \times 10^{-4}$). Significant associations were also seen in the analysis of quantity of cannabis use for 9 of the 10 SNP cutoffs, with the top association seen for those SNPs with $P \leqslant 0.05$ for schizophrenia ($R^2 = 0.85\%$, $P = 0.003$). No association was seen with age at initiation of use, although the association with quantity of use remained significant when number of years of usage was accounted for (results not shown).

As a secondary analysis, polygenic risk score for schizophrenia risk alleles with $P \leqslant 0.01$ (the threshold with the greatest association in the primary analysis) was examined within 990 twin pairs (608 dizygotic and 382 monozygotic) where data on cannabis use of both twins was available. Taking the mean polygenic risk score within each twin pair, an ordinal regression was performed to predict whether neither ($n = 272$), one ($n = 273$) or both twins ($n = 445$) were cannabis users. After correcting for age, sex and zygosity, a significant association was observed ($P = 0.001$). Those twin pairs where both reported using cannabis had the greatest burden of schizophrenia risk alleles, pairs with only one user were found to have an intermediate level and the

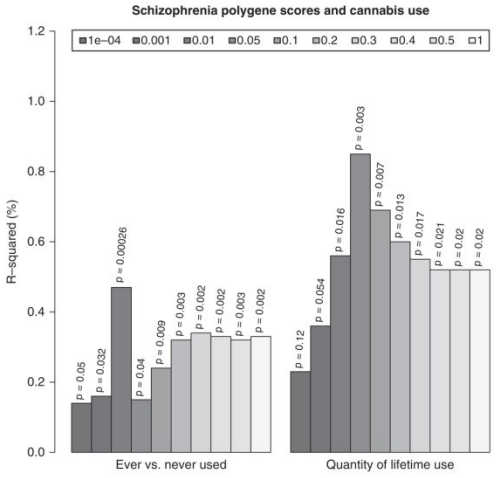


Figure 1. Results of polygenic risk scores for schizophrenia predicting variance explained (R^2) in cannabis use as both a binary trait of ever versus never, and as a quantitative trait of lifetime use within only users. Polygenic scores were created using different cutoffs for the inclusion of risk variants for schizophrenia, ranging from $P = 0.0001$ to 1.0.

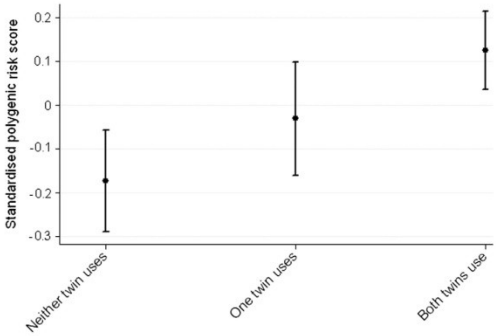


Figure 2. Mean standardized schizophrenia polygenic risk scores for pairs of twins when neither ($n = 272$), one ($n = 273$) or both twins ($n = 445$) had reported use of cannabis. An ordinal regression reported a significant association ($P = 0.001$).

lowest burden was found in pairs where neither twin reported use (see Figure 2).

DISCUSSION

Our results show that to some extent the association between cannabis and schizophrenia is due to a shared genetic aetiology across common variants. They suggest that individuals with an increased genetic predisposition to schizophrenia are both more likely to use cannabis and to use it in greater quantities. This is not to say that there is no causal relationship between use of cannabis and risk of schizophrenia, but it does establish that at least part of the association may be due to causal relationship in the opposite direction. Although the variance in cannabis use explained by schizophrenia polygenic risk scores is small, it is in line with other cross-phenotype analyses, largely due to the polygenic risk scores for schizophrenia predicting only ~7% of the variation for schizophrenia itself. Previous associations between polygenic risk scores for schizophrenia and other psychiatric illnesses, such as bipolar disorder, major depression and autism,[39] have shown effects of similar sizes. Further research will be needed to see whether the genetic overlap observed here is specific to cannabis use or is present across illicit drug use and addiction phenotypes, data for which was not widely available in this sample. For now, these findings have important implications for the current perception of cannabis use as a risk factor for schizophrenia, and other psychotic disorders.

However, it is worth noting that this association, if true, does not rule out the possibility of cannabis independently being a risk factor for schizophrenia. A bidirectional association between cannabis use and psychosis has previously been suggested.[40] Further, one caveat to interpreting the direction of causation concerns the discovery sample used to identify schizophrenia risk alleles. The schizophrenia GWAS sample will likely include many more cannabis users among cases than controls. This may lead to an excess of causal SNPs associated with cannabis use, as opposed to schizophrenia itself, identified as schizophrenia risk alleles. Only if the discovery schizophrenia sample was comprised entirely of non-cannabis users could causation be inferred without any risk of confounding. This is an important consideration as to whether polygenic risk scores overestimate individuals' un-modifiable genetic risk by including their genetic predisposition to modifiable environmental risk factors.

These results highlight the blurring between behavioural phenotypes and environment, and have wider implications for how we perceive supposedly environmental risks for disease. Individuals select their own environments based on their innate and learned preferences, and have their environments react to their own behaviour. Further, parents pass down both genes and environment to their children. All of these can contribute to gene–environment correlation, particularly with respect to behavioural traits. Several studies have shown that supposedly environmental risk factors such as urbanicity, religiosity and stressful life events have heritable components to them.[41–43] The existence of heritability for supposedly environmental risk factors does not mean they are inevitable, only that causality is more complicated to discern. Future studies will need to explore the matching of cases and controls on environmental risk variants to fully disentangle causation. This can be supplemented exploring the generation of polygenic risk scores for environmental risk factors, and their role in predicting disease status. The wider availability of genetic data in richly phenotyped samples should allow for the integration of genetics into an epidemiological framework, and so the discovery of gene–environment correlations where they exist.

With ongoing debate over the legalization of cannabis and the potential health risks it poses, understanding the association between its use and schizophrenia is a priority. It has previously been suggested that, even assuming an entirely causal relationship, the required reduction in the number of cannabis users to prevent one case of schizophrenia is in the thousands.[44] Our findings here highlight the possibility that this association might be bidirectional in causation, and that the risks of cannabis use could be overestimated. This is an important subtlety to consider when calculating the economic and health impact of cannabis use.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Robert Power was funded by the Medical Research Council and the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. This work was supported by National Institutes of Health Grants AA07535, AA0758O, AA07728, AA10249, AA13320, AA13321, AA14041, AA11998, AA17688, DA012854, DA018267, DA018660, DA23668 and DA019951; by Grants from the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498, 6136022, 628911 and 1047956); by Grants from the Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016 and DP0343921); and by the 5th Framework Programme (FP-5) GenomEUtwin Project (QLG2-CT-2002-01254). This research was further supported by the Centre for Research Excellence on Suicide Prevention (CRESP—Australia). KJHV is supported by the Netherlands Organization for Health Research and Development, ZonMW 31160212. We thank Richard Parker, Soad Hancock, Judith Moir, Sally Rodda, Pieta-Maree Shertock, Heather Park, Jill Wood, Pam Barton, Fran Husband, Adele Somerville, Ann Eldridge, Marlene Grace, Kerrie McAloney, Lisa Bowdler, Alexandre Todorov, Steven Crooks, David Smyth, Harry Beeby and Daniel Park. Finally, we thank the twins and their families for their participation.

REFERENCES

1 United Nations Office on Drugs and Crime. *World Drug Report 2013*. United Nations: New York, NY, USA, 2013.
2 Green B, Young R, Kavanagh D. Cannabis use and misuse prevalence among people with psychosis. *Br J Psychiatry* 2005; **187**: 306–313.
3 McGuire PK, Jones P, Harvey I, Bebbington P, Toone B, Lewis S *et al*. Cannabis and acute psychosis. *Schizophr Res* 1994; **13**: 161–167.
4 Thornicroft G. Cannabis and psychosis. Is there epidemiological evidence for an association?. *Br J Psychiatry* 1990; **157**: 25–33.
5 Tien AY, Anthony JC. Epidemiological analysis of alcohol and drug use as risk factors for psychotic experiences. *J Nerv Ment Dis* 1990; **178**: 473–480.
6 D'Souza DC, Perry E, MacDougall L, Ammerman Y, Cooper T, Wu YT *et al*. The psychotomimetic effects of intravenous delta-9-tetrahydrocannabinol in healthy

individuals: implications for psychosis. *Neuropsychopharmacology* 2004; **29**: 1558–1572.
7 Morrison PD, Zois V, McKeown DA, Lee TD, Holt DW, Powell JF *et al.* The acute effects of synthetic intravenous Delta9-tetrahydrocannabinol on psychosis, mood and cognitive functioning. *Psychol Med* 2009; **39**: 1607–1616.
8 van Os J, Bak M, Hanssen M, Bijl RV, de Graaf R, Verdoux H. Cannabis use and psychosis: a longitudinal population-based study. *Am J Epidemiol* 2002; **156**: 319–327.
9 Regier DA, Farmer ME, Rae DS, Locke BZ, Keith SJ, Judd LL *et al.* Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA* 1990; **264**: 2511–2518.
10 Arseneault L, Cannon M, Witton J, Murray RM. Causal association between cannabis and psychosis: examination of the evidence. *Br J Psychiatry* 2004; **184**: 110–117.
11 McGrath J, Welham J, Scott J, Varghese D, Degenhardt L, Hayatbakhsh MR *et al.* Association between cannabis use and psychosis-related outcomes using sibling pair analysis in a cohort of young adults. *Arch Gen Psychiatry* 2010; **67**: 440–447.
12 Zammit S, Allebeck P, Andreasson S, Lundberg I, Lewis G. Self reported cannabis use as a risk factor for schizophrenia in Swedish conscripts of 1969: historical cohort study. *BMJ* 2002; **325**: 1199.
13 Henquet C, Krabbendam L, Spauwen J, Kaplan C, Lieb R, Wittchen HU *et al.* Prospective cohort study of cannabis use, predisposition for psychosis, and psychotic symptoms in young people. *BMJ* 2005; **330**: 11.
14 Moore TH, Zammit S, Lingford-Hughes A, Barnes TR, Jones PB, Burke M *et al.* Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. *Lancet* 2007; **370**: 319–328.
15 Arseneault L, Cannon M, Poulton R, Murray R, Caspi A, Moffitt TE. Cannabis use in adolescence and risk for adult psychosis: longitudinal prospective study. *BMJ* 2002; **325**: 1212–1213.
16 Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M *et al.* Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380**: 2163–2196.
17 Addington J, Duchak V. Reasons for substance use in schizophrenia. *Acta Psychiatr Scand* 1997; **96**: 329–333.
18 Spencer C, Castle D, Michie PT. Motivations that maintain substance use among individuals with psychotic disorders. *Schizophr Bull* 2002; **28**: 233–247.
19 Ferdinand RF, Sondeijker F, van der Ende J, Selten JP, Huizink A, Verhulst FC. Cannabis use predicts future psychotic symptoms, and vice versa. *Addiction* 2005; **100**: 612–618.
20 Fergusson DM, Horwood LJ, Ridder EM. Tests of causal linkages between cannabis use and psychotic symptoms. *Addiction* 2005; **100**: 354–366.
21 Compton MT, Furman AC, Kaslow NJ. Lower negative symptom scores among cannabis-dependent patients with schizophrenia-spectrum disorders: preliminary evidence from an African American first-episode sample. *Schizophr Res* 2004; **71**: 61–64.
22 Peralta V, Cuesta MJ. Influence of cannabis abuse on schizophrenic psychopathology. *Acta Psychiatr Scand* 1992; **85**: 127–130.
23 Dixon L, Haas G, Weiden PJ, Sweeney J, Frances AJ. Drug abuse in schizophrenic patients: clinical correlates and reasons for use. *Am J Psychiatry* 1991; **148**: 224–230.
24 Henquet C, Di Forti M, Morrison P, Kuepper R, Murray RM. Gene-environment interplay between cannabis and psychosis. *Schizophr Bull* 2008; **34**: 1111–1121.
25 Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait—Evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 2003; **60**: 1187–1192.
26 Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA *et al.* Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
27 Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* 2013; **45**: 1150–1159.
28 Kendler KS, Schmitt E, Aggen SH, Prescott CA. Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Arch Gen Psychiatry* 2008; **65**: 674–682.
29 Agrawal A, Lynskey MT. The genetic epidemiology of cannabis use, abuse and dependence. *Addiction* 2006; **101**: 801–812.
30 Verweij KJ, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG *et al.* Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. *Addiction* 2010; **105**: 417–430.
31 Verweij KJ, Vinkhuyzen AA, Benyamin B, Lynskey MT, Quaye L, Agrawal A *et al.* The genetic aetiology of cannabis use initiation: a meta-analysis of genome-wide association studies and a SNP-based heritability estimation. *Addict Biol* 2013; **18**: 846–850.
32 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
33 Bucholz KK, Heath AC, Madden PA. Transitions in drinking in adolescent females: evidence from the Missouri adolescent female twin study. *Alcohol Clin Exp Res* 2000; **24**: 914–923.
34 Knopik VS, Heath AC, Madden PA, Bucholz KK, Slutske WS, Nelson EC *et al.* Genetic effects on alcohol dependence risk: re-evaluating the importance of psychiatric and other heritable risk factors. *Psychol Med* 2004; **34**: 1519–1530.
35 Heath AC, Whitfield JB, Martin NG, Pergadia ML, Goate AM, Lind PA *et al.* A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry* 2011; **70**: 513–518.
36 Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G *et al.* Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet* 2009; **85**: 750–755.
37 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
38 StataCorp. Stata Statistical Software: Release 12. StataCorp LP: College Station, TX, USA, 2011.
39 Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI *et al.* Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.
40 Foti DJ, Kotov R, Guey LT, Bromet EJ. Cannabis use and the course of schizophrenia: 10-year follow-up after first hospitalization. *Am J Psychiatry* 2010; **167**: 987–993.
41 Power RA, Wingenbach T, Cohen-Woods S, Uher R, Ng MY, Butler AW *et al.* Estimating the heritability of reporting stressful life events captured by common genetic variants. *Psychol Med* 2012; **43**: 1965–1971.
42 Whitfield JB, Zhu G, Heath AC, Martin NG. Choice of residential location: chance, family influences, or genes? *Twin Res Hum Genet* 2005; **8**: 22–26.
43 Winter T, Kaprio J, Viken RJ, Karvonen S, Rose RJ. Individual differences in adolescent religiosity in Finland: familial effects are modified by sex and region of residence. *Twin Res* 1999; **2**: 108–114.
44 Hickman M, Vickerman P, Macleod J, Lewis G, Zammit S, Kirkbride J *et al.* If cannabis caused schizophrenia--how many cannabis users may need to be prevented in order to prevent one case of schizophrenia? England and Wales calculations. *Addiction* 2009; **104**: 1856–1861.

8. Discussion

8.1 Summary

Psychiatric disorders reflect a wide range of illnesses that, while showing substantial aetiological and diagnostic overlap, vary greatly in prevalence, symptoms, onset, and heritability. The work presented in this thesis began by looking at the impact of these disorders on reproductive fitness, to understand if they are likely to undergo the same selection pressures and so have similar underlying genetic architectures. This identified that depression was not under negative selection, a finding that was then explored through molecular data. These results suggested the potential role of gene-environment interactions as an adaptive mechanism for variants contributing risk to depression, and were followed up by studies focusing on identifying such interactions. The conclusion of the thesis was though no gene-environmental interaction could be found to explain how the risk variants for major depression avoided negative selection, this may be due to substantial gene-environment correlation in the reporting and experiencing of environmental risk factors in psychiatric disorders which would confound such analyses. A more detailed discussion follows of the findings and how they relate to each other.

8.2 Overview of Findings

The first paper of the thesis examined epidemiological data from Sweden for evidence of the selection on psychiatric traits [Power and others 2013b]. The most striking finding of this analysis was that affected males showed a much greater reduction in fecundity than affected females, suggesting that psychiatric illness impacts the fitness of men to a greater extent and there may be sex-specific selection on the causal genetic variants. On an individual disorder basis, autism and schizophrenia showed the expected large reduction in fecundity within patients suggesting that genetic variants that predispose to risk are under negative selection and are maintained by *de novo* mutations. Anorexia, substance abuse and bipolar disorder showed a moderate reduction in fecundity, suggesting the role of *de novo mutations* may be less important. This supports a

distinction between bipolar disorder and schizophrenia in the literature of molecular studies of these disorders [Sullivan and others 2012], with a greater role of novel mutations observed in schizophrenia. Surprisingly, only a very small reduction was seen in depression patients which was further restricted to affected males not females. Even more surprisingly this reduction was more than compensated by the observed increase in fecundity in the siblings of affected individuals, suggesting that having certain combinations of the risk variants and environments for depression could be beneficial to an individual's fitness. This finding was based on modern levels of fecundity which may be confounded by a number of factors, such as the effects of hospitalisation and treatment, modern stigma of the disorders, and the use of contraceptives, which may give estimates of selection on these disorders that do not relate to historical levels throughout human evolution. This is an important distinction as the genetic architecture of traits is defined by selection over previous generations, rather than selection in the current generation. In the second paper of this thesis, it was supported by findings from a molecular study of 9 studies of major depression looking for an association with runs of homozygosity (ROH) [Power and others 2014a],which can be used to infer the impact of selection on a trait over previous generations. Inbreeding, as captured by ROH, exposes recessive variants, and an excess of recessive variants in association with a trait suggests that it has previously experienced negative selection. This is due to the fact that selection rapidly removes dominant deleterious mutations but both cannot as easily act on recessive variants and acts to select dominant advantageous variants. ROH were not associated with major depression status across a meta-analysis of nine studies, suggesting overall there was no evidence for negative selection on depression over previous generations.

However individual studies within this meta-analysis showed significant associations with ROH and significant heterogeneity across studies was observed, seemingly driven in part by the country the study took place in. This result and the findings from the epidemiological study, as well as previous research from others [Caspi and others 2003], suggested the potential role of differing

environmental backgrounds determining the impact of variants predisposing to MDD on fecundity. To examine this idea further, two studies were performed within the RADIANT consortium where access to detailed environment and genotype data was available. In the third paper of this thesis the interaction between childhood and adult environmental adversity was examined [Power and others 2013c], testing the mismatch hypothesis of depression whereby individuals are primed as children for a stressful or non-stressful environment in adulthood and the disorder manifests when this expectation is not met [Nederhof and Schmidt 2012] . This hypothesis also puts forward that some individuals show plasticity and can be 'programmed' by early experiences to develop behaviours that will maximise future gain (e.g. calibrating their level of anxiety to avoid harm but not be overly conservative), while other individuals show no plasticity but are more robust to environmental stressors. In this model of depression, the risk variants for depression would not always confer risk but rather are likely 'plasticity' markers of an individual's predisposition for priming in childhood or potentially their predisposition to misjudge the level of stress they should be primed for. However no evidence for the mismatch hypothesis was found in the interaction of environmental stressors across life or with a potential plasticity genetic risk variant (5-*HTTPLR*). This may have been due to the severity of the environmental stressors used, particularly the level of childhood maltreatment which was often severe, or the very improbable likelihood of success from using a candidate gene over an established replicated variant (which are currently lacking for depression) [Ripke and others 2013b]. The fourth paper of this thesis focused on a novel approach to indentifying novel variants for major depression not by looking for their interaction with environmental stressors, but by matching cases and controls on environment [Power and others 2013a]. In this analysis we observed no such variants, likely due to the large reduction in sample size that resulted from including only cases and controls with overlapping levels of stressful life events and the large discrepancy in their distributions in the two groups. In fact we observed a noticeable deflation in the expected number of associations when cases and controls were matched

on their reported number of stressful life events, suggesting that even adding more samples was unlikely to increase our chances of identifying novel genetic variants.

This deflation of associations may have been due to overcorrecting if the genetics of depression correlates with the reporting of stressful life events, or the genetics of reporting or experiencing such events. Stressful life events are themselves are known to be heritable [Billig and others 1996; Kendler and others 1993; Plomin and others 1990], as their likelihood of occurring is to some extent the result of an individual's personality and actions. In the last two papers of thesis we explored this in two ways. The fifth paper of the thesis focused on looking at the heritability of stressful life events, or at least the heritability of reporting such events, making use of genome-wide heritability estimation methods to calculate the heritability of a trait directly captured by molecular data rather than from an adoption or twin design [Power and others 2013d]. Such methods work by using genetic data to estimate the very low levels of relatedness between each pair of 'unrelated' individuals within the sample, as estimating to what extent this relatedness explained variation in phenotype. This found substantial heritability in stressful life events captured by the common variants observed, that appeared to be driven primarily by an individual's level of neuroticism. This itself is a heritable trait and suspected to overlap with depression in aetiology and definition, suggesting that those who are predisposed to being more depressed report more stressful life events. This may reflect a difference in perception of one's environment that may precede the onset of the disorder and confound the analysis of triggers of depression, and highlights the importance of having rigorously defined and recorded environmental measures due to the possibility of unreliable reporting. The sixth and last paper of this thesis showed that this gene-environment correlation is not restricted to the reporting of stressful events but also observed for traits such as drug use, showing that healthy individuals in the general population with a greater genetic predisposition to schizophrenia as defined by polygenic scoring are more likely to use cannabis [Power and others 2014b]. Polygenic scoring is a method that takes the findings from a genome-

wide association study to define which variants are more common in cases than controls (regardless of significance) and then counting them up in each individual of a independent sample. This provides each individual with a polygenic risk score which can be used to predict the same trait or other traits to test for genetic overlap, in this case taking the results from the most recent genome-wide association study of schizophrenia [Ripke and others 2013a] to test their cumulative association with cannabis use. This showed an association with both whether an individual had ever used cannabis and within users their quantity of use, but not with age at first use which had previously also been suggested to influence risk of psychosis. Given the well established association between cannabis use and risk of schizophrenia from epidemiological work [Arseneault and others 2004; McGrath and others 2010], this gene-environment correlations are important for how we identify and define environmental risk factors for psychiatric disorders. That is not to say that every gene-environment correlation removes the risk from environment, in fact the risk could still be mediated by the environment whereby individuals are genetically predisposed to seek out environments that are deleterious. Distinguishing situations where genes predispose individuals pleiotropically to both a disease and an environment through independent pathways (and so preventing the environment does not reduce disease risk), from situations where genes predispose individuals to an environment that increases the risk of a disease (and so mediates the pathway to risk) will be crucial for understanding when environmental interventions will yield success or not. Understanding how to integrate such gene-environment correlations into the analyses of gene-environment interactions will be crucial to developing methods to better identify the latter, and understand how genes and environment come together to predispose individuals to risk of a disorder.

8.3 Future directions

There are several potential areas of research that expand upon this work. Primarily is the wider application of polygenic methods, such as polygenic scoring and genome-wide heritability

estimates such as those available from GCTA [Purcell and others 2009; Yang and others 2011], to gene-environment interaction studies. These allow researchers to side step the main problem of such studies to date, that they must rely either on unreliable candidate genes that often fail to replicate or on genome-wide association data that allows the identification of novel genetic variants but with much reduced power.

These methods can also continue to be applied to the search for gene-environment correlations as in this thesis, such as following up the findings between schizophrenia and cannabis by looking at other illicit drugs or addictions. However to do so without any risk of confounding, greater information on the environmental exposures will be needed across datasets. For instance in the analysis of schizophrenia and cannabis use, the polygenic scores for schizophrenia were based on a genome-wide association study of schizophrenia that was blind to cannabis use. In this scenario the association with cannabis use could be due to cannabis use risk variants being picked up in the schizophrenia GWAS, as cases within this study are much more likely to also be users than controls. As such the GWAS of schizophrenia, with more cases with environmental risks for schizophrenia than controls, would pick up not just genetic variants that predispose directly to schizophrenia but also any variants that predispose to environments that increase risk to schizophrenia (assuming there is a heritable component to experiencing these environments). Ideally a follow up study could be performed where the discovery GWAS of schizophrenia had cases and controls matched on cannabis use, or the GWAS restricted to non-users in both groups. This would give a more definitive understanding of how the two traits' genetics overlap. It may also give us an estimate of how much of the heritability of schizophrenia comes from genetic predisposition to deleterious environments, and so how much of the heritability is actually modifiable.

Turning away from the gene-environment section of the thesis and back to the evolutionary analyses of depression, there are implications from finding no negative selection on depression. The

effect size and frequency of 'risk' variants should be curtailed by the selection against them in the population, meaning a trait with weaker selection against it should on average have variants of larger effect and so be more easily identified. However in the case of depression this is clearly not the case, given the difficulty in robustly identifying any risk variants to date [Ripke and others 2013b]. This could in part be due to the fact that the lack of selection is due to variants being beneficial in some circumstances (as seen in the analysis of fecundity of siblings of those with depression), meaning that many seemingly healthy controls will be carriers and reducing power to detect associations.

How exactly these variants might be beneficial is currently unknown, though some possibilities are apparent. Most psychiatric disorders and diseases can be viewed as a quantitative trait. In depression there is the extreme of low mood, pessimism and social withdrawal, though it is easy to imagine how an overly optimistic individual could also suffer reduced fitness from misjudging real risks and engaging in situations or individuals that were better avoided. Given the supposed association with stressful life events and loss of social support, depression may to some extent reflect an adaptive mechanism (or at least an adaptive mechanism malfunctioning) evolved to avoid risks when least able to deal with them. It is also worth noting that it has long been believed that psychiatric disorders, particularly psychosis, may overlap in families with beneficial traits. A large family studies of several hundred thousand in Sweden found that individuals with bipolar disorder and healthy siblings of people with schizophrenia and bipolar disorder are overrepresented in creative professions but those with major depression and their relatives are not [Kyaga and others 2011]. The results of this study provide support for the notion that creativity and bipolar disorder may co-segregate in families, however did not provide matching evidence for major depression.

Whether such associations with beneficial traits or gene-environments will provide novel insights into the genetic architecture of psychiatric disorders in the future is unclear. However the

increasing existence of large population datasets with access to detailed phenotypic, environmental and genomics data should allow future research to explore these questions and hopefully improve our understanding of the aetiologies of these disorders.

9. References

2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447(7145):661-678.

2010. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet 42(5):441-447.

Agrawal A, Lynskey MT. 2006. The genetic epidemiology of cannabis use, abuse and dependence. Addiction 101(6):801-812.

Andreasen NC. 1987. Creativity and mental illness: prevalence rates in writers and their first-degree relatives. Am J Psychiatry 144(10):1288-1292.

Angold A, Costello EJ, Worthman CM. 1998. Puberty and depression: the roles of age, pubertal status and pubertal timing. Psychol Med 28(1):51-61.

Angst F, Stassen HH, Clayton PJ, Angst J. 2002. Mortality of patients with mood disorders: follow-up over 34-38 years. J Affect Disord 68(2-3):167-181.

APA. 1994. Diagnostic and Statistical Manual of Mental Disorders Washington, DC: American Psychiatric Association

Arseneault L, Cannon M, Witton J, Murray RM. 2004. Causal association between cannabis and psychosis: examination of the evidence. Br J Psychiatry 184:110-117.

Barkley RA. 2008. Global issues related to the impact of untreated attention-deficit/hyperactivity disorder from childhood to young adulthood. Postgraduate Medicine 120(3):48-59.

Barth J, Schumacher M, Herrmann-Lingen C. 2004. Depression as a risk factor for mortality in patients with coronary heart disease: a meta-analysis. Psychosom Med 66(6):802-813.

Bebbington P, Ramana R. 1995. The epidemiology of bipolar affective disorder. Soc Psychiatry Psychiatr Epidemiol 30(6):279-292.

Becker JB, Hu M. 2008. Sex differences in drug abuse. Front Neuroendocrinol 29(1):36-47.

Bertram L, Tanzi RE. 2008. Thirty years of Alzheimer's disease genetics: the implications of systematic meta-analyses. Nat Rev Neurosci 9(10):768-778.

Betancur C. 2011. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. Brain Res 1380:42-77.

Bierut LJ, Goate AM, Breslau N, Johnson EO, Bertelsen S, Fox L, Agrawal A, Bucholz KK, Grucza R, Hesselbrock V, Kramer J, Kuperman S, Nurnberger J, Porjesz B, Saccone NL, Schuckit M, Tischfield J, Wang JC, Foroud T, Rice JP, Edenberg HJ. 2012. ADH1B is associated with alcohol dependence and alcohol consumption in populations of European and African ancestry. Mol Psychiatry 17(4):445-450.

Billig JP, Hershberger SL, Iacono WG, McGue M. 1996. Life events and personality in late adolescence: genetic and environmental relations. Behav Genet 26(6):543-554.

Boomsma DI, Saviouk V, Hottenga JJ, Distel MA, de Moor MH, Vink JM, Geels LM, van Beek JH, Bartels M, de Geus EJ, Willemsen G. 2010. Genetic epidemiology of attention deficit hyperactivity disorder (ADHD index) in adults. Plos One 5(5):e10621.

Boraska V, Franklin CS, Floyd JA, Thornton LM, Huckins LM, Southam L, Rayner NW, Tachmazidou I, Klump KL, Treasure J, Lewis CM, Schmidt U, Tozzi F, Kiezebrink K, Hebebrand J, Gorwood P, Adan RA, Kas MJ, Favaro A, Santonastaso P, Fernandez-Aranda F, Gratacos M, Rybakowski F, Dmitrzak-Weglarz M, Kaprio J, Keski-Rahkonen A, Raevuori A, Van Furth EF, Slof-Op 't Landt MC, Hudson JI, Reichborn-Kjennerud T, Knudsen GP, Monteleone P, Kaplan AS, Karwautz A, Hakonarson H, Berrettini WH, Guo Y, Li D, Schork NJ, Komaki G, Ando T, Inoko H, Esko T, Fischer K, Mannik K, Metspalu A, Baker JH, Cone RD, Dackor J, Desocio JE, Hilliard CE, O'Toole JK, Pantel J, Szatkiewicz JP, Taico C, Zerwas S, Trace SE, Davis OS, Helder S, Buhren K, Burghardt R, de

Zwaan M, Egberts K, Ehrlich S, Herpertz-Dahlmann B, Herzog W, Imgart H, Scherag A, Scherag S, Zipfel S, Boni C, Ramoz N, Versini A, Brandys MK, Danner UN, de Kovel C, Hendriks J, Koeleman BP, Ophoff RA, Strengman E, van Elburg AA, Bruson A, Clementi M, Degortes D, Forzan M, Tenconi E, Docampo E, Escaramis G, Jimenez-Murcia S, Lissowska J, Rajewski A, Szeszenia-Dabrowska N, Slopien A, Hauser J, Karhunen L, Meulenbelt I, Slagboom PE, Tortorella A, Maj M, Dedoussis G, Dikeos D, Gonidakis F, Tziouvas K, Tsitsika A, Papezova H, Slachtova L, Martaskova D, Kennedy JL, Levitan RD, Yilmaz Z, Huemer J, Koubek D, Merl E, Wagner G, Lichtenstein P, Breen G, Cohen-Woods S, Farmer A, McGuffin P, Cichon S, Giegling I, Herms S, Rujescu D, Schreiber S, Wichmann HE, Dina C, Sladek R, Gambaro G, Soranzo N, Julia A, Marsal S, Rabionet R, Gaborieau V, Dick DM, Palotie A, Ripatti S, Widen E, Andreassen OA, Espeseth T, Lundervold A, Reinvang I, Steen VM, Le Hellard S, Mattingsdal M, Ntalla I, Bencko V, Foretova L, Janout V, Navratilova M, Gallinger S, Pinto D, Scherer SW, Aschauer H, Carlberg L, Schosser A, Alfredsson L, Ding B, Klareskog L, Padyukov L, Courtet P, Guillaume S, Jaussent I, Finan C, Kalsi G, Roberts M, Logan DW, Peltonen L, Ritchie GR, Barrett JC, Estivill X, Hinney A, Sullivan PF, Collier DA, Zeggini E, Bulik CM. 2014. A genome-wide association study of anorexia nervosa. Mol Psychiatry.

Breslau N, Johnson EO, Hiripi E, Kessler R. 2001. Nicotine dependence in the United States: prevalence, trends, and smoking persistence. Arch Gen Psychiatry 58(9):810-816.

Brown AS, Schaefer CA, Wyatt RJ, Begg MD, Goetz R, Bresnahan MA, Harkavy-Friedman J, Gorman JM, Malaspina D, Susser ES. 2002. Paternal age and risk of schizophrenia in adult offspring. Am J Psychiatry 159(9):1528-1533.

Bulik CM, Slof-Op't Landt MC, van Furth EF, Sullivan PF. 2007. The genetics of anorexia nervosa. Annu Rev Nutr 27:263-275.

Bundy H, Stahl D, MacCabe JH. 2011. A systematic review and meta-analysis of the fertility of patients with schizophrenia and their unaffected relatives. Acta Psychiat Scand 123(2):98-106.

Byrne EM, Gehrman PR, Medland SE, Nyholt DR, Heath AC, Madden PA, Hickie IB, Van Duijn CM, Henders AK, Montgomery GW, Martin NG, Wray NR. 2013. A genome-wide association study of sleep habits and insomnia. Am J Med Genet B Neuropsychiatr Genet 162B(5):439-451.

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R. 2003. Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. Science 301(5631):386-389.

Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA. 1993. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261(5123):921-923.

D'Onofrio BM, Rickert ME, Frans E, Kuja-Halkola R, Almqvist C, Sjolander A, Larsson H, Lichtenstein P. 2014. Paternal age at childbearing and offspring psychiatric and academic morbidity. JAMA Psychiatry 71(4):432-438.

Davis LK, Yu D, Keenan CL, Gamazon ER, Konkashbaev AI, Derks EM, Neale BM, Yang J, Lee SH, Evans P, Barr CL, Bellodi L, Benarroch F, Berrio GB, Bienvenu OJ, Bloch MH, Blom RM, Bruun RD, Budman CL, Camarena B, Campbell D, Cappi C, Cardona Silgado JC, Cath DC, Cavallini MC, Chavira DA, Chouinard S, Conti DV, Cook EH, Coric V, Cullen BA, Deforce D, Delorme R, Dion Y, Edlund CK, Egberts K, Falkai P, Fernandez TV, Gallagher PJ, Garrido H, Geller D, Girard SL, Grabe HJ, Grados MA, Greenberg BD, Gross-Tsur V, Haddad S, Heiman GA, Hemmings SM, Hounie AG, Illmann C, Jankovic J, Jenike MA, Kennedy JL, King RA, Kremeyer B, Kurlan R, Lanzagorta N, Leboyer M, Leckman JF, Lennertz L, Liu C, Lochner C, Lowe TL, Macciardi F, McCracken JT, McGrath LM, Mesa Restrepo SC, Moessner R, Morgan J, Muller H, Murphy DL, Naarden AL, Ochoa WC, Ophoff RA, Osiecki L, Pakstis AJ, Pato MT, Pato CN, Piacentini J, Pittenger C, Pollak Y, Rauch SL, Renner TJ, Reus VI, Richter MA, Riddle MA, Robertson MM, Romero R, Rosario MC, Rosenberg D, Rouleau GA, Ruhrmann S, Ruiz-Linares A, Sampaio AS, Samuels J, Sandor P, Sheppard B, Singer

HS, Smit JH, Stein DJ, Strengman E, Tischfield JA, Valencia Duarte AV, Vallada H, Van Nieuwerburgh F, Veenstra-Vanderweele J, Walitza S, Wang Y, Wendland JR, Westenberg HG, Shugart YY, Miguel EC, McMahon W, Wagner M, Nicolini H, Posthuma D, Hanna GL, Heutink P, Denys D, Arnold PD, Oostra BA, Nestadt G, Freimer NB, Pauls DL, Wray NR, Stewart SE, Mathews CA, Knowles JA, Cox NJ, Scharf JM. 2013. Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. PLoS Genet 9(10):e1003864.

de Candia TR, Lee SH, Yang J, Browning BL, Gejman PV, Levinson DF, Mowry BJ, Hewitt JK, Goddard ME, O'Donovan MC, Purcell SM, Posthuma D, Visscher PM, Wray NR, Keller MC. 2013. Additive genetic variation in schizophrenia risk is shared by populations of African and European descent. Am J Hum Genet 93(3):463-470.

Dudbridge F, Gusnanto A. 2008. Estimation of significance thresholds for genomewide association scans. Genet Epidemiol 32(3):227-234.

Durkin MS, Maenner MJ, Newschaffer CJ, Lee LC, Cunniff CM, Daniels JL, Kirby RS, Leavitt L, Miller L, Zahorodny W, Schieve LA. 2008. Advanced parental age and the risk of autism spectrum disorder. Am J Epidemiol 168(11):1268-1276.

Faraone SV, Chen WJ, Goldstein JM, Tsuang MT. 1994. Gender differences in age at onset of schizophrenia. Br J Psychiatry 164(5):625-629.

Faraone SV, Perlis RH, Doyle AE, Smoller JW, Goralnick JJ, Holmgren MA, Sklar P. 2005. Molecular genetics of attention-deficit/hyperactivity disorder. Biol Psychiatry 57(11):1313-1323.

Folstein SE, Rosen-Sheidley B. 2001. Genetics of autism: complex aetiology for a heterogeneous disorder. Nat Rev Genet 2(12):943-955.

Fombonne E. 2005. Epidemiology of autistic disorder and other pervasive developmental disorders. J Clin Psychiatry 66 Suppl 10:3-8.

Frans EM, McGrath JJ, Sandin S, Lichtenstein P, Reichenberg A, Langstrom N, Hultman CM. 2011. Advanced paternal and grandpaternal age and schizophrenia: a three-generation perspective. Schizophr Res 133(1-3):120-124.

Frans EM, Sandin S, Reichenberg A, Langstrom N, Lichtenstein P, McGrath JJ, Hultman CM. 2013. Autism risk across generations: a population-based study of advancing grandpaternal and paternal age. JAMA Psychiatry 70(5):516-521.

Frans EM, Sandin S, Reichenberg A, Lichtenstein P, Langstrom N, Hultman CM. 2008. Advancing paternal age and bipolar disorder. Arch Gen Psychiat 65(9):1034-1040.

Freeman EW, Sammel MD, Liu L, Gracia CR, Nelson DB, Hollander L. 2004. Hormones and menopausal status as predictors of depression in women in transition to menopause. Arch Gen Psychiatry 61(1):62-70.

Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, Georgieva L, Rees E, Palta P, Ruderfer DM, Carrera N, Humphreys I, Johnson JS, Roussos P, Barker DD, Banks E, Milanova V, Grant SG, Hannon E, Rose SA, Chambert K, Mahajan M, Scolnick EM, Moran JL, Kirov G, Palotie A, McCarroll SA, Holmans P, Sklar P, Owen MJ, Purcell SM, O'Donovan MC. 2014. De novo mutations in schizophrenia implicate synaptic networks. Nature 506(7487):179-184.

Gale CR, Batty GD, McIntosh AM, Porteous DJ, Deary IJ, Rasmussen F. 2013. Is bipolar disorder more common in highly intelligent people? A cohort study of a million men. Mol Psychiatry 18(2):190-194.

Gamsiz ED, Viscidi EW, Frederick AM, Nagpal S, Sanders SJ, Murtha MT, Schmidt M, Triche EW, Geschwind DH, State MW, Istrail S, Cook EH, Jr., Devlin B, Morrow EM. 2013. Intellectual disability is associated with increased runs of homozygosity in simplex autism. Am J Hum Genet 93(1):103-109.

Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL. 2006. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry 63(2):168-174.

Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH, Anton R, Preuss UW, Ridinger M, Rujescu D, Wodarz N, Zill P, Zhao H, Farrer LA. 2014. Genome-wide association study of alcohol dependence:significant findings in African- and European-Americans including novel risk loci. Mol Psychiatry 19(1):41-49.

Goriely A, McGrath JJ, Hultman CM, Wilkie AO, Malaspina D. 2013. "Selfish spermatogonial selection": a novel mechanism for the association between advanced paternal age and neurodevelopmental disorders. Am J Psychiatry 170(6):599-608.

Greenwald BS, Kramer-Ginsberg E, Marin DB, Laitman LB, Hermann CK, Mohs RC, Davis KL. 1989. Dementia with coexistent major depression. Am J Psychiatry 146(11):1472-1478.

Grozeva D, Kirov G, Ivanov D, Jones IR, Jones L, Green EK, St Clair DM, Young AH, Ferrier N, Farmer AE, McGuffin P, Holmans PA, Owen MJ, O'Donovan MC, Craddock N. 2010. Rare copy number variants: a point of rarity in genetic risk for bipolar disorder and schizophrenia. Arch Gen Psychiatry 67(4):318-327.

Guerreiro RJ, Lohmann E, Kinsella E, Bras JM, Luu N, Gurunlian N, Dursun B, Bilgic B, Santana I, Hanagasi H, Gurvit H, Gibbs JR, Oliveira C, Emre M, Singleton A. 2012. Exome sequencing reveals an unexpected genetic cause of disease: NOTCH3 mutation in a Turkish family with Alzheimer's disease. Neurobiol Aging 33(5):1008 e1017-1023.

Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schurmann B, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frolich L, Hampel H, Hull M, Rujescu D, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Muhleisen TW, Nothen MM, Moebus S, Jockel KH, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J. 2009. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet 41(10):1088-1093.

Hasin DS, Goodwin RD, Stinson FS, Grant BF. 2005. Epidemiology of major depressive disorder - Results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiat 62(10):1097-1106.

Hasin DS, Stinson FS, Ogburn E, Grant BF. 2007. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry 64(7):830-842.

Herbert PS, Jr. 1959. Creativity and mental illness: a study of 60 creative patients who needed hospitalization. The Psychiatric quarterly 33:534-547.

Hoek HW. 2006. Incidence, prevalence and mortality of anorexia nervosa and other eating disorders. Curr Opin Psychiatry 19(4):389-394.

Hoek HW, van Hoeken D. 2003. Review of the prevalence and incidence of eating disorders. Int J Eat Disord 34(4):383-396.

Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, Abraham R, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Jones N, Stretton A, Thomas C, Richards A, Ivanov D, Widdowson C, Chapman J, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Beaumont H, Warden D, Wilcock G, Love S, Kehoe PG, Hooper NM, Vardy ER,

Hardy J, Mead S, Fox NC, Rossor M, Collinge J, Maier W, Jessen F, Ruther E, Schurmann B, Heun R, Kolsch H, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frolich L, Hampel H, Gallacher J, Hull M, Rujescu D, Giegling I, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Muhleisen TW, Nothen MM, Moebus S, Jockel KH, Klopp N, Wichmann HE, Pankratz VS, Sando SB, Aasly JO, Barcikowska M, Wszolek ZK, Dickson DW, Graff-Radford NR, Petersen RC, van Duijn CM, Breteler MM, Ikram MA, DeStefano AL, Fitzpatrick AL, Lopez O, Launer LJ, Seshadri S, Berr C, Campion D, Epelbaum J, Dartigues JF, Tzourio C, Alperovitch A, Lathrop M, Feulner TM, Friedrich P, Riehle C, Krawczak M, Schreiber S, Mayhaus M, Nicolhaus S, Wagenpfeil S, Steinberg S, Stefansson H, Stefansson K, Snaedal J, Bjornsson S, Jonsson PV, Chouraki V, Genier-Boley B, Hiltunen M, Soininen H, Combarros O, Zelenika D, Delepine M, Bullido MJ, Pasquier F, Mateo I, Frank-Garcia A, Porcellini E, Hanon O, Coto E, Alvarez V, Bosco P, Siciliano G, Mancuso M, Panza F, Solfrizzi V, Nacmias B, Sorbi S, Bossu P, Piccardi P, Arosio B, Annoni G, Seripa D, Pilotto A, Scarpini E, Galimberti D, Brice A, Hannequin D, Licastro F, Jones L, Holmans PA, Jonsson T, Riemenschneider M, Morgan K, Younkin SG, Owen MJ, O'Donovan M, Amouyel P, Williams J. 2011. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat Genet 43(5):429-435.

Hosang GM, Korszun A, Jones L, Jones I, McGuffin P, Farmer AE. 2012. Life-event specificity: bipolar disorder compared with unipolar depression. Br J Psychiatry 201:458-465.

Hudson JI, Hiripi E, Pope HG, Jr., Kessler RC. 2007. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. Biol Psychiatry 61(3):348-358.

Hultman CM, Sandin S, Levine SZ, Lichtenstein P, Reichenberg A. 2011. Advancing paternal age and risk of autism: new evidence from a population-based study and a meta-analysis of epidemiological studies. Mol Psychiatry 16(12):1203-1212.

Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, Yamrom B, Lee YH, Narzisi G, Leotta A, Kendall J, Grabowska E, Ma B, Marks S, Rodgers L, Stepansky A, Troge J, Andrews P, Bekritsky M, Pradhan K, Ghiban E, Kramer M, Parla J, Demeter R, Fulton LL, Fulton RS, Magrini VJ, Ye K, Darnell JC, Darnell RB, Mardis ER, Wilson RK, Schatz MC, McCombie WR, Wigler M. 2012. De novo gene disruptions in children on the autistic spectrum. Neuron 74(2):285-299.

Jamison KR. 1989. Mood disorders and patterns of creativity in British writers and artists. Psychiatry 52(2):125-134.

Jensen AR. 1983. Effects of inbreeding on mental-ability factors. Personality and Individual Differences 4(1):71-87.

Joukamaa M, Heliovaara M, Knekt P, Aromaa A, Raitasalo R, Lehtinen V. 2001. Mental disorders and cause-specific mortality. Br J Psychiatry 179:498-502.

Juda A. 1949. The relationship between highest mental capacity and psychic abnormalities. Am J Psychiatry 106(4):296-307.

Keller M, Miller G. 2006. Resolving the evolutionary paradox of common, harmful, heritable mental disorders. Am J Med Genet B 141B(7):719-719.

Keller MC, Simonson MA, Ripke S, Neale BM, Gejman PV, Howrigan DP, Lee SH, Lencz T, Levinson DF, Sullivan PF. 2012. Runs of homozygosity implicate autozygosity as a schizophrenia risk factor. PLoS Genet 8(4):e1002656.

Keller MC, Visscher PM, Goddard ME. 2011. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics 189(1):237-249.

Kendler KS, Neale M, Kessler R, Heath A, Eaves L. 1993. A twin study of recent life events and difficulties. Arch Gen Psychiatry 50(10):789-796.

Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS. 2003. The epidemiology of major depressive disorder - Results from the National Comorbidity Survey Replication (NCS-R). Jama-J Am Med Assoc 289(23):3095-3105.

Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D, Moran J, Chambert K, Toncheva D, Georgieva L, Grozeva D, Fjodorova M, Wollerton R, Rees E, Nikolov I, van de Lagemaat LN, Bayes A, Fernandez E, Olason PI, Bottcher Y, Komiyama NH, Collins MO, Choudhary J, Stefansson K, Stefansson H, Grant SG, Purcell S, Sklar P, O'Donovan MC, Owen MJ. 2012. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Mol Psychiatry 17(2):142-153.

Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA, Sigurdsson A, Jonasdottir A, Wong WS, Sigurdsson G, Walters GB, Steinberg S, Helgason H, Thorleifsson G, Gudbjartsson DF, Helgason A, Magnusson OT, Thorsteinsdottir U, Stefansson K. Rate of de novo mutations and the importance of father's age to disease risk. Nature 488(7412):471-475.

Kuntsi J, Eley TC, Taylor A, Hughes C, Asherson P, Caspi A, Moffitt TE. 2004. Co-occurrence of ADHD and low IQ has genetic origins. Am J Med Genet B Neuropsychiatr Genet 124B(1):41-47.

Kyaga S, Landen M, Boman M, Hultman CM, Langstrom N, Lichtenstein P. 2013. Mental illness, suicide and creativity: 40-year prospective total population study. J Psychiatr Res 47(1):83-90.

Kyaga S, Lichtenstein P, Boman M, Hultman C, Langstrom N, Landen M. 2011. Creativity and mental disorder: family study of 300,000 people with severe mental disorder. Br J Psychiatry 199:373-379.

Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, Combarros O, Zelenika D, Bullido MJ, Tavernier B, Letenneur L, Bettens K, Berr C, Pasquier F, Fievet N, Barberger-Gateau P, Engelborghs S, De Deyn P, Mateo I, Franck A, Helisalmi S, Porcellini E, Hanon O, de Pancorbo MM, Lendon C, Dufouil C, Jaillard C, Leveillard T, Alvarez V, Bosco P, Mancuso M, Panza F, Nacmias B, Bossu P, Piccardi P, Annoni G, Seripa D, Galimberti D, Hannequin D, Licastro F, Soininen H, Ritchie K, Blanche H, Dartigues JF, Tzourio C, Gut I, Van Broeckhoven C, Alperovitch A, Lathrop M, Amouyel P. 2009. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet 41(10):1094-1099.

Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thorton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Moron FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fievet N, Huentelman MW, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuiness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossu P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Deniz Naranjo MC, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannefelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH, Jr., Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RF, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JS, Boerwinkle E, Riemenschneider M, Boada M, Hiltuenen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nothen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA,

van Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P. 2013. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 45(12):1452-1458.

Larsson H, Chang Z, D'Onofrio BM, Lichtenstein P. 2013. The heritability of clinically diagnosed attention deficit hyperactivity disorder across the lifespan. Psychol Med:1-7.

Lee SH, Harold D, Nyholt DR, Goddard ME, Zondervan KT, Williams J, Montgomery GW, Wray NR, Visscher PM. 2013a. Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. Hum Mol Genet 22(4):832-841.

Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, Mowry BJ, Thapar A, Goddard ME, Witte JS, Absher D, Agartz I, Akil H, Amin F, Andreassen OA, Anjorin A, Anney R, Anttila V, Arking DE, Asherson P, Azevedo MH, Backlund L, Badner JA, Bailey AJ, Banaschewski T, Barchas JD, Barnes MR, Barrett TB, Bass N, Battaglia A, Bauer M, Bayes M, Bellivier F, Bergen SE, Berrettini W, Betancur C, Bettecken T, Biederman J, Binder EB, Black DW, Blackwood DH, Bloss CS, Boehnke M, Boomsma DI, Breen G, Breuer R, Bruggeman R, Cormican P, Buccola NG, Buitelaar JK, Bunney WE, Buxbaum JD, Byerley WF, Byrne EM, Caesar S, Cahn W, Cantor RM, Casas M, Chakravarti A, Chambert K, Choudhury K, Cichon S, Cloninger CR, Collier DA, Cook EH, Coon H, Cormand B, Corvin A, Coryell WH, Craig DW, Craig IW, Crosbie J, Cuccaro ML, Curtis D, Czamara D, Datta S, Dawson G, Day R, De Geus EJ, Degenhardt F, Djurovic S, Donohoe GJ, Doyle AE, Duan J, Dudbridge F, Duketis E, Ebstein RP, Edenberg HJ, Elia J, Ennis S, Etain B, Fanous A, Farmer AE, Ferrier IN, Flickinger M, Fombonne E, Foroud T, Frank J, Franke B, Fraser C, Freedman R, Freimer NB, Freitag CM, Friedl M, Frisen L, Gallagher L, Gejman PV, Georgieva L, Gershon ES, Geschwind DH, Giegling I, Gill M, Gordon SD, Gordon-Smith K, Green EK, Greenwood TA, Grice DE, Gross M, Grozeva D, Guan W, Gurling H, De Haan L, Haines JL, Hakonarson H, Hallmayer J, Hamilton SP, Hamshere ML, Hansen TF, Hartmann AM, Hautzinger M, Heath AC, Henders AK, Herms S, Hickie IB, Hipolito M, Hoefels S, Holmans PA, Holsboer F, Hoogendijk WJ, Hottenga JJ, Hultman CM, Hus V, Ingason A, Ising M, Jamain S, Jones EG, Jones I, Jones L, Tzeng JY, Kahler AK, Kahn RS, Kandaswamy R, Keller MC, Kennedy JL, Kenny E, Kent L, Kim Y, Kirov GK, Klauck SM, Klei L, Knowles JA, Kohli MA, Koller DL, Konte B, Korszun A, Krabbendam L, Krasucki R, Kuntsi J, Kwan P, Landen M, Langstrom N, Lathrop M, Lawrence J, Lawson WB, Leboyer M, Ledbetter DH, Lee PH, Lencz T, Lesch KP, Levinson DF, Lewis CM, Li J, Lichtenstein P, Lieberman JA, Lin DY, Linszen DH, Liu C, Lohoff FW, Loo SK, Lord C, Lowe JK, Lucae S, MacIntyre DJ, Madden PA, Maestrini E, Magnusson PK, Mahon PB, Maier W, Malhotra AK, Mane SM, Martin CL, Martin NG, Mattheisen M, Matthews K, Mattingsdal M, McCarroll SA, McGhee KA, McGough JJ, McGrath PJ, McGuffin P, McInnis MG, McIntosh A, McKinney R, McLean AW, McMahon FJ, McMahon WM, McQuillin A, Medeiros H, Medland SE, Meier S, Melle I, Meng F, Meyer J, Middeldorp CM, Middleton L, Milanova V, Miranda A, Monaco AP, Montgomery GW, Moran JL, Moreno-De-Luca D, Morken G, Morris DW, Morrow EM, Moskvina V, Muglia P, Muhleisen TW, Muir WJ, Muller-Myhsok B, Murtha M, Myers RM, Myin-Germeys I, Neale MC, Nelson SF, Nievergelt CM, Nikolov I, Nimgaonkar V, Nolen WA, Nothen MM, Nurnberger JI, Nwulia EA, Nyholt DR, O'Dushlaine C, Oades RD, Olincy A, Oliveira G, Olsen L, Ophoff RA, Osby U, Owen MJ, Palotie A, Parr JR, Paterson AD, Pato CN, Pato MT, Penninx BW, Pergadia ML, Pericak-Vance MA, Pickard BS, Pimm J, Piven J, Posthuma D, Potash JB, Poustka F, Propping P, Puri V, Quested DJ, Quinn EM, Ramos-Quiroga JA, Rasmussen HB, Raychaudhuri S, Rehnstrom K, Reif A, Ribases M, Rice JP, Rietschel M, Roeder K, Roeyers H, Rossin L, Rothenberger A, Rouleau G, Ruderfer D, Rujescu D, Sanders AR, Sanders SJ, Santangelo SL, Sergeant JA, Schachar R, Schalling M, Schatzberg AF, Scheftner WA, Schellenberg GD, Scherer SW, Schork NJ, Schulze TG, Schumacher J, Schwarz M, Scolnick E, Scott LJ, Shi J, Shilling PD, Shyn SI, Silverman JM, Slager SL, Smalley SL, Smit JH, Smith EN,

Sonuga-Barke EJ, St Clair D, State M, Steffens M, Steinhausen HC, Strauss JS, Strohmaier J, Stroup TS, Sutcliffe JS, Szatmari P, Szelinger S, Thirumalai S, Thompson RC, Todorov AA, Tozzi F, Treutlein J, Uhr M, van den Oord EJ, Van Grootheest G, Van Os J, Vicente AM, Vieland VJ, Vincent JB, Visscher PM, Walsh CA, Wassink TH, Watson SJ, Weissman MM, Werge T, Wienker TF, Wijsman EM, Willemsen G, Williams N, Willsey AJ, Witt SH, Xu W, Young AH, Yu TW, Zammit S, Zandi PP, Zhang P, Zitman FG, Zollner S, Devlin B, Kelsoe JR, Sklar P, Daly MJ, O'Donovan MC, Craddock N, Sullivan PF, Smoller JW, Kendler KS, Wray NR. 2013b. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet 45(9):984-994.

Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. 2012. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics 28(19):2540-2542.

Levinson DF. 2006. The genetics of depression: a review. Biol Psychiatry 60(2):84-92.

Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I, Williams NM, Schwab SG, Pulver AE, Faraone SV, Brzustowicz LM, Kaufmann CA, Garver DL, Gurling HM, Lindholm E, Coon H, Moises HW, Byerley W, Shaw SH, Mesen A, Sherrington R, O'Neill FA, Walsh D, Kendler KS, Ekelund J, Paunio T, Lonnqvist J, Peltonen L, O'Donovan MC, Owen MJ, Wildenauer DB, Maier W, Nestadt G, Blouin JL, Antonarakis SE, Mowry BJ, Silverman JM, Crowe RR, Cloninger CR, Tsuang MT, Malaspina D, Harkavy-Friedman JM, Svrakic DM, Bassett AS, Holcomb J, Kalsi G, McQuillin A, Brynjolfson J, Sigmundsson T, Petursson H, Jazin E, Zoega T, Helgason T. 2003. Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. Am J Hum Genet 73(1):34-48.

Lichtenstein P, Bjork C, Hultman CM, Scolnick E, Sklar P, Sullivan PF. 2006. Recurrence risks for schizophrenia in a Swedish national cohort. Psychol Med 36(10):1417-1425.

Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, Hultman CM. 2009. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. Lancet 373(9659):234-239.

Liu JZ, Tozzi F, Waterworth DM, Pillai SG, Muglia P, Middleton L, Berrettini W, Knouff CW, Yuan X, Waeber G, Vollenweider P, Preisig M, Wareham NJ, Zhao JH, Loos RJ, Barroso I, Khaw KT, Grundy S, Barter P, Mahley R, Kesaniemi A, McPherson R, Vincent JB, Strauss J, Kennedy JL, Farmer A, McGuffin P, Day R, Matthews K, Bakke P, Gulsvik A, Lucae S, Ising M, Brueckl T, Horstmann S, Wichmann HE, Rawal R, Dahmen N, Lamina C, Polasek O, Zgaga L, Huffman J, Campbell S, Kooner J, Chambers JC, Burnett MS, Devaney JM, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Epstein S, Wilson JF, Wild SH, Campbell H, Vitart V, Reilly MP, Li M, Qu L, Wilensky R, Matthai W, Hakonarson HH, Rader DJ, Franke A, Wittig M, Schafer A, Uda M, Terracciano A, Xiao X, Busonero F, Scheet P, Schlessinger D, St Clair D, Rujescu D, Abecasis GR, Grabe HJ, Teumer A, Volzke H, Petersmann A, John U, Rudan I, Hayward C, Wright AF, Kolcic I, Wright BJ, Thompson JR, Balmforth AJ, Hall AS, Samani NJ, Anderson CA, Ahmad T, Mathew CG, Parkes M, Satsangi J, Caulfield M, Munroe PB, Farrall M, Dominiczak A, Worthington J, Thomson W, Eyre S, Barton A, Mooser V, Francks C, Marchini J. 2010. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. Nat Genet 42(5):436-440.

Lobo A, Launer LJ, Fratiglioni L, Andersen K, Di Carlo A, Breteler MM, Copeland JR, Dartigues JF, Jagger C, Martinez-Lage J, Soininen H, Hofman A. 2000. Prevalence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. Neurology 54(11 Suppl 5):S4-9.

Lubke GH, Laurin C, Amin N, Hottenga JJ, Willemsen G, van Grootheest G, Abdellaoui A, Karssen LC, Oostra BA, van Duijn CM, Penninx BW, Boomsma DI. 2013. Genome-wide analyses of borderline personality features. Mol Psychiatry.

Lucas AR, Beard CM, O'Fallon WM, Kurland LT. 1991. 50-year trends in the incidence of anorexia nervosa in Rochester, Minn.: a population-based study. Am J Psychiatry 148(7):917-922.
Ludwig AM. 1992. Creative achievement and psychopathology: comparison among professions. American journal of psychotherapy 46(3):330-356.
MacCabe JH, Koupil I, Leon DA. 2009. Lifetime reproductive output over two generations in patients with psychosis and their unaffected siblings: the Uppsala 1915-1929 Birth Cohort Multigenerational Study. Psychol Med 39(10):1667-1676.
MacCabe JH, Lambe MP, Cnattingius S, Sham PC, David AS, Reichenberg A, Murray RM, Hultman CM. 2010. Excellent school performance at age 16 and risk of adult bipolar disorder: national cohort study. Br J Psychiatry 196(2):109-115.
Malaspina D, Harlap S, Fennig S, Heiman D, Nahon D, Feldman D, Susser ES. 2001. Advancing paternal age and the risk of schizophrenia. Arch Gen Psychiat 58(4):361-367.
Malhotra D, McCarthy S, Michaelson JJ, Vacic V, Burdick KE, Yoon S, Cichon S, Corvin A, Gary S, Gershon ES, Gill M, Karayiorgou M, Kelsoe JR, Krastoshevsky O, Krause V, Leibenluft E, Levy DL, Makarov V, Bhandari A, Malhotra AK, McMahon FJ, Nothen MM, Potash JB, Rietschel M, Schulze TG, Sebat J. 2011. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. Neuron 72(6):951-963.
Mansour H, Kandil K, Wood J, Fathi W, Elassy M, Ibrahim I, Salah H, Yassin A, Elsayed H, Tobar S, El-Boraie H, Eissa A, Elhadidy M, Ibrahim NE, El-Bahaei W, Nimgaonkar VL. 2011. Reduced Fertility and Fecundity among Patients with Bipolar I Disorder and Schizophrenia in Egypt. Psychiatry Investig 8(3):214-220.
McGrath J, Saha S, Chant D, Welham J. 2008. Schizophrenia: a concise overview of incidence, prevalence, and mortality. Epidemiologic reviews 30:67-76.
McGrath J, Welham J, Scott J, Varghese D, Degenhardt L, Hayatbakhsh MR, Alati R, Williams GM, Bor W, Najman JM. 2010. Association between cannabis use and psychosis-related outcomes using sibling pair analysis in a cohort of young adults. Arch Gen Psychiatry 67(5):440-447.
McGrath JJ, Hearle J, Jenner L, Plant K, Drummond A, Barkla JM. 1999. The fertility and fecundity of patients with psychoses. Acta Psychiatr Scand 99(6):441-446.
McGrath JJ, Petersen L, Agerbo E, Mors O, Mortensen PB, Pedersen CB. 2014. A comprehensive assessment of parental age and psychiatric disorders. JAMA Psychiatry 71(3):301-309.
McGuffin P. 1979. Is schizophrenia an HLA-associated disease? Psychol Med 9(4):721-728.
McGuffin P, Katz R, Watkins S, Rutherford J. 1996. A hospital-based twin register of the heritability of DSM-IV unipolar depression. Arch Gen Psychiat 53(2):129-136.
McNaughton D, Knight W, Guerreiro R, Ryan N, Lowe J, Poulter M, Nicholl DJ, Hardy J, Revesz T, Rossor M, Collinge J, Mead S. 2012. Duplication of amyloid precursor protein (APP), but not prion protein (PRNP) gene is a significant cause of early onset dementia in a large UK series. Neurobiol Aging 33(2):426 e413-421.
McQuillin A, Bass N, Anjorin A, Lawrence J, Kandaswamy R, Lydall G, Moran J, Sklar P, Purcell S, Gurling H. 2011. Analysis of genetic deletions and duplications in the University College London bipolar disorder case control sample. Eur J Hum Genet 19(5):588-592.
Morton NE. 1978. Effect of inbreeding on IQ and mental retardation. Proc Natl Acad Sci U S A 75(8):3906-3908.
Mueller TI, Leon AC, Keller MB, Solomon DA, Endicott J, Coryell W, Warshaw M, Maser JD. 1999. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. Am J Psychiatry 156(7):1000-1006.
Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J, Gallins PJ, Buxbaum JD, Jarvik GP, Crane PK, Larson EB, Bird TD, Boeve BF, Graff-Radford NR, De Jager PL, Evans D, Schneider JA, Carrasquillo MM, Ertekin-Taner N, Younkin SG, Cruchaga C, Kauwe JS, Nowotny P, Kramer P,

Hardy J, Huentelman MJ, Myers AJ, Barmada MM, Demirci FY, Baldwin CT, Green RC, Rogaeva E, St George-Hyslop P, Arnold SE, Barber R, Beach T, Bigio EH, Bowen JD, Boxer A, Burke JR, Cairns NJ, Carlson CS, Carney RM, Carroll SL, Chui HC, Clark DG, Corneveaux J, Cotman CW, Cummings JL, DeCarli C, DeKosky ST, Diaz-Arrastia R, Dick M, Dickson DW, Ellis WG, Faber KM, Fallon KB, Farlow MR, Ferris S, Frosch MP, Galasko DR, Ganguli M, Gearing M, Geschwind DH, Ghetti B, Gilbert JR, Gilman S, Giordani B, Glass JD, Growdon JH, Hamilton RL, Harrell LE, Head E, Honig LS, Hulette CM, Hyman BT, Jicha GA, Jin LW, Johnson N, Karlawish J, Karydas A, Kaye JA, Kim R, Koo EH, Kowall NW, Lah JJ, Levey AI, Lieberman AP, Lopez OL, Mack WJ, Marson DC, Martiniuk F, Mash DC, Masliah E, McCormick WC, McCurry SM, McDavid AN, McKee AC, Mesulam M, Miller BL, Miller CA, Miller JW, Parisi JE, Perl DP, Peskind E, Petersen RC, Poon WW, Quinn JF, Rajbhandary RA, Raskind M, Reisberg B, Ringman JM, Roberson ED, Rosenberg RN, Sano M, Schneider LS, Seeley W, Shelanski ML, Slifer MA, Smith CD, Sonnen JA, Spina S, Stern RA, Tanzi RE, Trojanowski JQ, Troncoso JC, Van Deerlin VM, Vinters HV, Vonsattel JP, Weintraub S, Welsh-Bohmer KA, Williamson J, Woltjer RL, Cantwell LB, Dombroski BA, Beekly D, Lunetta KL, Martin ER, Kamboh MI, Saykin AJ, Reiman EM, Bennett DA, Morris JC, Montine TJ, Goate AM, Blacker D, Tsuang DW, Hakonarson H, Kukull WA, Foroud TM, Haines JL, Mayeux R, Pericak-Vance MA, Farrer LA, Schellenberg GD. 2011. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet 43(5):436-441.

Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V, Polak P, Yoon S, Maguire J, Crawford EL, Campbell NG, Geller ET, Valladares O, Schafer C, Liu H, Zhao T, Cai G, Lihm J, Dannenfelser R, Jabado O, Peralta Z, Nagaswamy U, Muzny D, Reid JG, Newsham I, Wu Y, Lewis L, Han Y, Voight BF, Lim E, Rossin E, Kirby A, Flannick J, Fromer M, Shakir K, Fennell T, Garimella K, Banks E, Poplin R, Gabriel S, DePristo M, Wimbish JR, Boone BE, Levy SE, Betancur C, Sunyaev S, Boerwinkle E, Buxbaum JD, Cook EH, Jr., Devlin B, Gibbs RA, Roeder K, Schellenberg GD, Sutcliffe JS, Daly MJ. 2012. Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 485(7397):242-245.

Neale BM, Medland SE, Ripke S, Asherson P, Franke B, Lesch KP, Faraone SV, Nguyen TT, Schafer H, Holmans P, Daly M, Steinhausen HC, Freitag C, Reif A, Renner TJ, Romanos M, Romanos J, Walitza S, Warnke A, Meyer J, Palmason H, Buitelaar J, Vasquez AA, Lambregts-Rommelse N, Gill M, Anney RJ, Langely K, O'Donovan M, Williams N, Owen M, Thapar A, Kent L, Sergeant J, Roeyers H, Mick E, Biederman J, Doyle A, Smalley S, Loo S, Hakonarson H, Elia J, Todorov A, Miranda A, Mulas F, Ebstein RP, Rothenberger A, Banaschewski T, Oades RD, Sonuga-Barke E, McGough J, Nisenbaum L, Middleton F, Hu X, Nelson S. 2010. Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 49(9):884-897.

Nederhof E, Schmidt MV. 2012. Mismatch or cumulative stress: Toward an integrated hypothesis of programming effects. Physiol Behav 106(5):691-700.

Ng MY, Levinson DF, Faraone SV, Suarez BK, DeLisi LE, Arinami T, Riley B, Paunio T, Pulver AE, Irmansyah, Holmans PA, Escamilla M, Wildenauer DB, Williams NM, Laurent C, Mowry BJ, Brzustowicz LM, Maziade M, Sklar P, Garver DL, Abecasis GR, Lerer B, Fallin MD, Gurling HM, Gejman PV, Lindholm E, Moises HW, Byerley W, Wijsman EM, Forabosco P, Tsuang MT, Hwu HG, Okazaki Y, Kendler KS, Wormley B, Fanous A, Walsh D, O'Neill FA, Peltonen L, Nestadt G, Lasseter VK, Liang KY, Papadimitriou GM, Dikeos DG, Schwab SG, Owen MJ, O'Donovan MC, Norton N, Hare E, Raventos H, Nicolini H, Albus M, Maier W, Nimgaonkar VL, Terenius L, Mallet J, Jay M, Godard S, Nertney D, Alexander M, Crowe RR, Silverman JM, Bassett AS, Roy MA, Merette C, Pato CN, Pato MT, Roos JL, Kohn Y, Amann-Zalcenstein D, Kalsi G, McQuillin A, Curtis D, Brynjolfson J, Sigmundsson T, Petursson H, Sanders AR, Duan J, Jazin E, Myles-Worsley M, Karayiorgou M,

Lewis CM. 2009. Meta-analysis of 32 genome-wide linkage studies of schizophrenia. Mol Psychiatry 14(8):774-785.
Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, Bustamante CD. 2008. Genes mirror geography within Europe. Nature 456(7218):98-101.
O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD, Turner EH, Stanaway IB, Vernot B, Malig M, Baker C, Reilly B, Akey JM, Borenstein E, Rieder MJ, Nickerson DA, Bernier R, Shendure J, Eichler EE. 2012. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature 485(7397):246-250.
Osby U, Brandt L, Correia N, Ekbom A, Sparen P. 2001. Excess mortality in bipolar and unipolar disorder in Sweden. Arch Gen Psychiat 58(9):844-850.
Pavlova B, Uher R, Dragomirecka E, Papezova H. 2009. Trends in hospital admissions for eating disorders in a country undergoing a socio-cultural transition, the Czech Republic 1981-2005. Soc Psychiatry Psychiatr Epidemiol 45(5):541-550.
Pedersen CB, McGrath J, Mortensen PB, Petersen L. 2014. The importance of father's age to schizophrenia risk. Mol Psychiatry 19(5):530.
Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Almeida J, Bacchelli E, Bader GD, Bailey AJ, Baird G, Battaglia A, Berney T, Bolshakova N, Bolte S, Bolton PF, Bourgeron T, Brennan S, Brian J, Bryson SE, Carson AR, Casallo G, Casey J, Chung BH, Cochrane L, Corsello C, Crawford EL, Crossett A, Cytrynbaum C, Dawson G, de Jonge M, Delorme R, Drmic I, Duketis E, Duque F, Estes A, Farrar P, Fernandez BA, Folstein SE, Fombonne E, Freitag CM, Gilbert J, Gillberg C, Glessner JT, Goldberg J, Green A, Green J, Guter SJ, Hakonarson H, Heron EA, Hill M, Holt R, Howe JL, Hughes G, Hus V, Igliozzi R, Kim C, Klauck SM, Kolevzon A, Korvatska O, Kustanovich V, Lajonchere CM, Lamb JA, Laskawiec M, Leboyer M, Le Couteur A, Leventhal BL, Lionel AC, Liu XQ, Lord C, Lotspeich L, Lund SC, Maestrini E, Mahoney W, Mantoulan C, Marshall CR, McConachie H, McDougle CJ, McGrath J, McMahon WM, Merikangas A, Migita O, Minshew NJ, Mirza GK, Munson J, Nelson SF, Noakes C, Noor A, Nygren G, Oliveira G, Papanikolaou K, Parr JR, Parrini B, Paton T, Pickles A, Pilorge M, Piven J, Ponting CP, Posey DJ, Poustka A, Poustka F, Prasad A, Ragoussis J, Renshaw K, Rickaby J, Roberts W, Roeder K, Roge B, Rutter ML, Bierut LJ, Rice JP, Salt J, Sansom K, Sato D, Segurado R, Sequeira AF, Senman L, Shah N, Sheffield VC, Soorya L, Sousa I, Stein O, Sykes N, Stoppioni V, Strawbridge C, Tancredi R, Tansey K, Thiruvahindrapduram B, Thompson AP, Thomson S, Tryfon A, Tsiantis J, Van Engeland H, Vincent JB, Volkmar F, Wallace S, Wang K, Wang Z, Wassink TH, Webber C, Weksberg R, Wing K, Wittemeyer K, Wood S, Wu J, Yaspan BL, Zurawiecki D, Zwaigenbaum L, Buxbaum JD, Cantor RM, Cook EH, Coon H, Cuccaro ML, Devlin B, Ennis S, Gallagher L, Geschwind DH, Gill M, Haines JL, Hallmayer J, Miller J, Monaco AP, Nurnberger JI, Jr., Paterson AD, Pericak-Vance MA, Schellenberg GD, Szatmari P, Vicente AM, Vieland VJ, Wijsman EM, Scherer SW, Sutcliffe JS, Betancur C. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. Nature 466(7304):368-372.
Plomin R, Lichtenstein P, Pedersen NL, McClearn GE, Nesselroade JR. 1990. Genetic influence on life events during the last half of the life span. Psychology and aging 5(1):25-30.
Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA. 2007. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. Am J Psychiatry 164(6):942-948.
Post F. 1994. Creativity and psychopathology. A study of 291 world-famous men. Br J Psychiatry 165(2):22-34.
Pottier C, Hannequin D, Coutant S, Rovelet-Lecrux A, Wallon D, Rousseau S, Legallic S, Paquet C, Bombois S, Pariente J, Thomas-Anterion C, Michon A, Croisile B, Etcharry-Bouyx F, Berr C, Dartigues JF, Amouyel P, Dauchel H, Boutoleau-Bretonniere C, Thauvin C, Frebourg T, Lambert

JC, Campion D. 2012. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. Mol Psychiatry 17(9):875-879.

Power RA, Cohen-Woods S, Ng MY, Butler AW, Craddock N, Korszun A, Jones L, Jones I, Gill M, Rice JP, Maier W, Zobel A, Mors O, Placentino A, Rietschel M, Aitchison KJ, Tozzi F, Muglia P, Breen G, Farmer AE, McGuffin P, Lewis CM, Uher R. 2013a. Genome-wide association analysis accounting for environmental factors through propensity-score matching: application to stressful live events in major depressive disorder. Am J Med Genet B Neuropsychiatr Genet 162B(6):521-529.

Power RA, Keller MC, Ripke S, Abdellaoui A, Wray NR, Sullivan PF, Breen G. 2014a. A recessive genetic model and runs of homozygosity in major depressive disorder. Am J Med Genet B Neuropsychiatr Genet 165B(2):157-166.

Power RA, Kyaga S, Uher R, MacCabe JH, Langstrom N, Landen M, McGuffin P, Lewis CM, Lichtenstein P, Svensson AC. 2013b. Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. JAMA Psychiatry 70(1):22-30.

Power RA, Lecky-Thompson L, Fisher HL, Cohen-Woods S, Hosang GM, Uher R, Powell-Smith G, Keers R, Tropeano M, Korszun A, Jones L, Jones I, Owen MJ, Craddock N, Craig IW, Farmer AE, McGuffin P. 2013c. The interaction between child maltreatment, adult stressful life events and the 5-HTTLPR in major depression. J Psychiatr Res 47(8):1032-1035.

Power RA, Verweij KJ, Zuhair M, Montgomery GW, Henders AK, Heath AC, Madden PA, Medland SE, Wray NR, Martin NG. 2014b. Genetic predisposition to schizophrenia associated with increased use of cannabis. Mol Psychiatry.

Power RA, Wingenbach T, Cohen-Woods S, Uher R, Ng MY, Butler AW, Ising M, Craddock N, Owen MJ, Korszun A, Jones L, Jones I, Gill M, Rice JP, Maier W, Zobel A, Mors O, Placentino A, Rietschel M, Lucae S, Holsboer F, Binder EB, Keers R, Tozzi F, Muglia P, Breen G, Craig IW, Muller-Myhsok B, Kennedy JL, Strauss J, Vincent JB, Lewis CM, Farmer AE, McGuffin P. 2013d. Estimating the heritability of reporting stressful life events captured by common genetic variants. Psychol Med 43(9):1965-1971.

Priebe L, Degenhardt FA, Herms S, Haenisch B, Mattheisen M, Nieratschker V, Weingarten M, Witt S, Breuer R, Paul T, Alblas M, Moebus S, Lathrop M, Leboyer M, Schreiber S, Grigoroiu-Serbanescu M, Maier W, Propping P, Rietschel M, Nothen MM, Cichon S, Muhleisen TW. 2012. Genome-wide survey implicates the influence of copy number variants (CNVs) in the development of early-onset bipolar disorder. Mol Psychiatry 17(4):421-432.

Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, O'Dushlaine C, Chambert K, Bergen SE, Kahler A, Duncan L, Stahl E, Genovese G, Fernandez E, Collins MO, Komiyama NH, Choudhary JS, Magnusson PK, Banks E, Shakir K, Garimella K, Fennell T, DePristo M, Grant SG, Haggarty SJ, Gabriel S, Scolnick EM, Lander ES, Hultman CM, Sullivan PF, McCarroll SA, Sklar P. 2014. A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506(7487):185-190.

Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P. 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460(7256):748-752.

Qiu C, Kivipelto M, von Strauss E. 2009. Epidemiology of Alzheimer's disease: occurrence, determinants, and strategies toward intervention. Dialogues Clin Neurosci 11(2):111-128.

Reichenberg A, Gross R, Weiser M, Bresnahan M, Silverman J, Harlap S, Rabinowitz J, Shulman C, Malaspina D, Lubin G, Knobler HY, Davidson M, Susser E. 2006. Advancing paternal age and autism. Arch Gen Psychiat 63(9):1026-1032.

Rietveld CA, Medland SE, Derringer J, Yang J, Esko T, Martin NW, Westra HJ, Shakhbazov K, Abdellaoui A, Agrawal A, Albrecht E, Alizadeh BZ, Amin N, Barnard J, Baumeister SE, Benke KS, Bielak LF, Boatman JA, Boyle PA, Davies G, de Leeuw C, Eklund N, Evans DS, Ferhmann R, Fischer K, Gieger

C, Gjessing HK, Hagg S, Harris JR, Hayward C, Holzapfel C, Ibrahim-Verbaas CA, Ingelsson E, Jacobsson B, Joshi PK, Jugessur A, Kaakinen M, Kanoni S, Karjalainen J, Kolcic I, Kristiansson K, Kutalik Z, Lahti J, Lee SH, Lin P, Lind PA, Liu Y, Lohman K, Loitfelder M, McMahon G, Vidal PM, Meirelles O, Milani L, Myhre R, Nuotio ML, Oldmeadow CJ, Petrovic KE, Peyrot WJ, Polasek O, Quaye L, Reinmaa E, Rice JP, Rizzi TS, Schmidt H, Schmidt R, Smith AV, Smith JA, Tanaka T, Terracciano A, van der Loos MJ, Vitart V, Volzke H, Wellmann J, Yu L, Zhao W, Allik J, Attia JR, Bandinelli S, Bastardot F, Beauchamp J, Bennett DA, Berger K, Bierut LJ, Boomsma DI, Bultmann U, Campbell H, Chabris CF, Cherkas L, Chung MK, Cucca F, de Andrade M, De Jager PL, De Neve JE, Deary IJ, Dedoussis GV, Deloukas P, Dimitriou M, Eiriksdottir G, Elderson MF, Eriksson JG, Evans DM, Faul JD, Ferrucci L, Garcia ME, Gronberg H, Guethnason V, Hall P, Harris JM, Harris TB, Hastie ND, Heath AC, Hernandez DG, Hoffmann W, Hofman A, Holle R, Holliday EG, Hottenga JJ, Iacono WG, Illig T, Jarvelin MR, Kahonen M, Kaprio J, Kirkpatrick RM, Kowgier M, Latvala A, Launer LJ, Lawlor DA, Lehtimaki T, Li J, Lichtenstein P, Lichtner P, Liewald DC, Madden PA, Magnusson PK, Makinen TE, Masala M, McGue M, Metspalu A, Mielck A, Miller MB, Montgomery GW, Mukherjee S, Nyholt DR, Oostra BA, Palmer LJ, Palotie A, Penninx BW, Perola M, Peyser PA, Preisig M, Raikkonen K, Raitakari OT, Realo A, Ring SM, Ripatti S, Rivadeneira F, Rudan I, Rustichini A, Salomaa V, Sarin AP, Schlessinger D, Scott RJ, Snieder H, St Pourcain B, Starr JM, Sul JH, Surakka I, Svento R, Teumer A, Tiemeier H, van Rooij FJ, Van Wagoner DR, Vartiainen E, Viikari J, Vollenweider P, Vonk JM, Waeber G, Weir DR, Wichmann HE, Widen E, Willemsen G, Wilson JF, Wright AF, Conley D, Davey-Smith G, Franke L, Groenen PJ, Johannesson M, Kardia SL, Krueger RF, Laibson D, Martin NG, Meyer MN, Posthuma D, Thurik AR, Timpson NJ, Uitterlinden AG, van Duijn CM, Visscher PM, Benjamin DJ, Cesarini D, Koellinger PD. 2013. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. Science 340(6139):1467-1471.

Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S, Bergen SE, Collins AL, Crowley JJ, Fromer M, Kim Y, Lee SH, Magnusson PK, Sanchez N, Stahl EA, Williams S, Wray NR, Xia K, Bettella F, Borglum AD, Bulik-Sullivan BK, Cormican P, Craddock N, de Leeuw C, Durmishi N, Gill M, Golimbet V, Hamshere ML, Holmans P, Hougaard DM, Kendler KS, Lin K, Morris DW, Mors O, Mortensen PB, Neale BM, O'Neill FA, Owen MJ, Milovancevic MP, Posthuma D, Powell J, Richards AL, Riley BP, Ruderfer D, Rujescu D, Sigurdsson E, Silagadze T, Smit AB, Stefansson H, Steinberg S, Suvisaari J, Tosato S, Verhage M, Walters JT, Levinson DF, Gejman PV, Laurent C, Mowry BJ, O'Donovan MC, Pulver AE, Schwab SG, Wildenauer DB, Dudbridge F, Shi J, Albus M, Alexander M, Campion D, Cohen D, Dikeos D, Duan J, Eichhammer P, Godard S, Hansen M, Lerer FB, Liang KY, Maier W, Mallet J, Nertney DA, Nestadt G, Norton N, Papadimitriou GN, Ribble R, Sanders AR, Silverman JM, Walsh D, Williams NM, Wormley B, Arranz MJ, Bakker S, Bender S, Bramon E, Collier D, Crespo-Facorro B, Hall J, Iyegbe C, Jablensky A, Kahn RS, Kalaydjieva L, Lawrie S, Lewis CM, Linszen DH, Mata I, McIntosh A, Murray RM, Ophoff RA, Van Os J, Walshe M, Weisbrod M, Wiersma D, Donnelly P, Barroso I, Blackwell JM, Brown MA, Casas JP, Corvin AP, Deloukas P, Duncanson A, Jankowski J, Markus HS, Mathew CG, Palmer CN, Plomin R, Rautanen A, Sawcer SJ, Trembath RC, Viswanathan AC, Wood NW, Spencer CC, Band G, Bellenguez C, Freeman C, Hellenthal G, Giannoulatou E, Pirinen M, Pearson RD, Strange A, Su Z, Vukcevic D, Langford C, Hunt SE, Edkins S, Gwilliam R, Blackburn H, Bumpstead SJ, Dronov S, Gillman M, Gray E, Hammond N, Jayakumar A, McCann OT, Liddle J, Potter SC, Ravindrarajah R, Ricketts M, Tashakkori-Ghanbaria A, Waller MJ, Weston P, Widaa S, Whittaker P, McCarthy MI, Stefansson K, Scolnick E, Purcell S, McCarroll SA, Sklar P, Hultman CM, Sullivan PF. 2013a. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet.

Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA, Lin DY, Duan J, Ophoff RA, Andreassen OA, Scolnick E, Cichon S, Clair DS, Corvin A, Gurling H, Werge T, Rujescu D,

Blackwood DHR, Pato CN, Malhotra AK, Purcell S, Dudbridge F, Neale BM, Rossin L, Visscher PM, Posthuma D, Ruderfer DM, Fanous A, Stefansson H, Steinberg S, Mowry BJ, Golimbet V, De Hert M, Jonsson EG, Bitter I, Pietilainen OPH, Collier DA, Tosato S, Agartz I, Albus M, Alexander M, Amdur RL, Amin F, Bass N, Bergen SE, Black DW, Borglum AD, Brown MA, Bruggeman R, Buccola NG, Byerley WF, Cahn W, Cantor RM, Carr VJ, Catts SV, Choudhury K, Cloninger CR, Cormican P, Craddock N, Danoy PA, Datta S, De Haan L, Demontis D, Dikeos D, Djurovic S, Donnelly P, Donohoe G, Duong L, Dwyer S, Fink-Jensen A, Freedman R, Freimer NB, Friedl M, Georgieva L, Giegling I, Gill M, Glenthoj B, Godard S, Hamshere M, Hansen M, Hansen T, Hartmann AM, Henskens FA, Hougaard DM, Hultman CM, Ingason A, Jablensky AV, Jakobsen KD, Jay M, Jurgens G, Kahn R, Keller MC, Kenis G, Kenny E, Kim Y, Kirov GK, Konnerth H, Konte B, Krabbendam L, Krasucki R, Lasseter VK, Laurent C, Lawrence J, Lencz T, Lerer FB, Liang KY, Lichtenstein P, Lieberman JA, Linszen DH, Lonnqvist J, Loughland CM, Maclean AW, Maher BS, Maier W, Mallet J, Malloy P, Mattheisen M, Mattingsdal M, McGhee KA, McGrath JJ, McIntosh A, McLean DE, McQuillin A, Melle I, Michie PT, Milanova V, Morris DW, Mors O, Mortensen PB, Moskvina V, Muglia P, Myin-Germeys I, Nertney DA, Nestadt G, Nielsen J, Nikolov I, Nordentoft M, Norton N, Nothen MM, O'Dushlaine CT, Olincy A, Olsen L, O'Neill FA, Orntoft TF, Owen MJ, Pantelis C, Papadimitriou G, Pato MT, Peltonen L, Petursson H, Pickard B, Pimm J, Pulver AE, Puri V, Quested D, Quinn EM, Rasmussen HB, Rethelyi JM, Ribble R, Rietschel M, Riley BP, Ruggeri M, Schall U, Schulze TG, Schwab SG, Scott RJ, Shi JX, Sigurdsson E, Silverman JM, Spencer CCA, Stefansson K, Strange A, Strengman E, Stroup TS, Suvisaari J, Terenius L, Thirumalai S, Thygesen JH, Timm S, Toncheva D, van den Oord E, van Os J, van Winkel R, Veldink J, Walsh D, Wang AG, Wiersma D, Wildenauer DB, Williams HJ, Williams NM, Wormley B, Zammit S, Sullivan PF, O'Donovan MC, Daly MJ, Gejman PV, Genome-Wide SP. 2011. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 43(10):969-U977.

Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G, Byrne EM, Blackwood DH, Boomsma DI, Cichon S, Heath AC, Holsboer F, Lucae S, Madden PA, Martin NG, McGuffin P, Muglia P, Noethen MM, Penninx BP, Pergadia ML, Potash JB, Rietschel M, Lin D, Muller-Myhsok B, Shi J, Steinberg S, Grabe HJ, Lichtenstein P, Magnusson P, Perlis RH, Preisig M, Smoller JW, Stefansson K, Uher R, Kutalik Z, Tansey KE, Teumer A, Viktorin A, Barnes MR, Bettecken T, Binder EB, Breuer R, Castro VM, Churchill SE, Coryell WH, Craddock N, Craig IW, Czamara D, De Geus EJ, Degenhardt F, Farmer AE, Fava M, Frank J, Gainer VS, Gallagher PJ, Gordon SD, Goryachev S, Gross M, Guipponi M, Henders AK, Herms S, Hickie IB, Hoefels S, Hoogendijk W, Hottenga JJ, Iosifescu DV, Ising M, Jones I, Jones L, Jung-Ying T, Knowles JA, Kohane IS, Kohli MA, Korszun A, Landen M, Lawson WB, Lewis G, Macintyre D, Maier W, Mattheisen M, McGrath PJ, McIntosh A, McLean A, Middeldorp CM, Middleton L, Montgomery GM, Murphy SN, Nauck M, Nolen WA, Nyholt DR, O'Donovan M, Oskarsson H, Pedersen N, Scheftner WA, Schulz A, Schulze TG, Shyn SI, Sigurdsson E, Slager SL, Smit JH, Stefansson H, Steffens M, Thorgeirsson T, Tozzi F, Treutlein J, Uhr M, van den Oord EJ, Van Grootheest G, Volzke H, Weilburg JB, Willemsen G, Zitman FG, Neale B, Daly M, Levinson DF, Sullivan PF. 2013b. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 18(4):497-511.

Robinson EB, Koenen KC, McCormick MC, Munir K, Hallett V, Happe F, Plomin R, Ronald A. Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). Arch Gen Psychiatry 68(11):1113-1121.

Ronald A, Hoekstra RA. 2011. Autism spectrum disorders and autistic traits: a decade of new twin studies. Am J Med Genet B Neuropsychiatr Genet 156B(3):255-274.

Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerriere A, Vital A, Dumanchin C, Feuillette S, Brice A, Vercelletto M, Dubas F, Frebourg T, Campion D. 2006. APP locus duplication causes

autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. Nat Genet 38(1):24-26.

Rucker JJ, Breen G, Pinto D, Pedroso I, Lewis CM, Cohen-Woods S, Uher R, Schosser A, Rivera M, Aitchison KJ, Craddock N, Owen MJ, Jones L, Jones I, Korszun A, Muglia P, Barnes MR, Preisig M, Mors O, Gill M, Maier W, Rice J, Rietschel M, Holsboer F, Farmer AE, Craig IW, Scherer SW, McGuffin P. 2013. Genome-wide association analysis of copy number variation in recurrent depressive disorder. Mol Psychiatry 18(2):183-189.

Rutter M. 2005. Incidence of autism spectrum disorders: changes over time and their meaning. Acta Paediatr 94(1):2-15.

Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, Mason CE, Bilguvar K, Celestino-Soper PB, Choi M, Crawford EL, Davis L, Wright NR, Dhodapkar RM, DiCola M, DiLullo NM, Fernandez TV, Fielding-Singh V, Fishman DO, Frahm S, Garagaloyan R, Goh GS, Kammela S, Klei L, Lowe JK, Lund SC, McGrew AD, Meyer KA, Moffat WJ, Murdoch JD, O'Roak BJ, Ober GT, Pottenger RS, Raubeson MJ, Song Y, Wang Q, Yaspan BL, Yu TW, Yurkiewicz IR, Beaudet AL, Cantor RM, Curland M, Grice DE, Gunel M, Lifton RP, Mane SM, Martin DM, Shaw CA, Sheldon M, Tischfield JA, Walsh CA, Morrow EM, Ledbetter DH, Fombonne E, Lord C, Martin CL, Brooks AI, Sutcliffe JS, Cook EH, Jr., Geschwind D, Roeder K, Devlin B, State MW. 2011. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron 70(5):863-885.

Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL, Walker MF, Ober GT, Teran NA, Song Y, El-Fishawy P, Murtha RC, Choi M, Overton JD, Bjornson RD, Carriero NJ, Meyer KA, Bilguvar K, Mane SM, Sestan N, Lifton RP, Gunel M, Roeder K, Geschwind DH, Devlin B, State MW. 2012. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485(7397):237-241.

Satin JR, Linden W, Phillips MJ. 2009. Depression as a predictor of disease progression and mortality in cancer patients: a meta-analysis. Cancer 115(22):5349-5361.

Scharf JM, Yu D, Mathews CA, Neale BM, Stewart SE, Fagerness JA, Evans P, Gamazon E, Edlund CK, Service SK, Tikhomirov A, Osiecki L, Illmann C, Pluzhnikov A, Konkashbaev A, Davis LK, Han B, Crane J, Moorjani P, Crenshaw AT, Parkin MA, Reus VI, Lowe TL, Rangel-Lugo M, Chouinard S, Dion Y, Girard S, Cath DC, Smit JH, King RA, Fernandez TV, Leckman JF, Kidd KK, Kidd JR, Pakstis AJ, State MW, Herrera LD, Romero R, Fournier E, Sandor P, Barr CL, Phan N, Gross-Tsur V, Benarroch F, Pollak Y, Budman CL, Bruun RD, Erenberg G, Naarden AL, Lee PC, Weiss N, Kremeyer B, Berrio GB, Campbell DD, Cardona Silgado JC, Ochoa WC, Mesa Restrepo SC, Muller H, Valencia Duarte AV, Lyon GJ, Leppert M, Morgan J, Weiss R, Grados MA, Anderson K, Davarya S, Singer H, Walkup J, Jankovic J, Tischfield JA, Heiman GA, Gilbert DL, Hoekstra PJ, Robertson MM, Kurlan R, Liu C, Gibbs JR, Singleton A, Hardy J, Strengman E, Ophoff RA, Wagner M, Moessner R, Mirel DB, Posthuma D, Sabatti C, Eskin E, Conti DV, Knowles JA, Ruiz-Linares A, Rouleau GA, Purcell S, Heutink P, Oostra BA, McMahon WM, Freimer NB, Cox NJ, Pauls DL. 2013. Genome-wide association study of Tourette's syndrome. Mol Psychiatry 18(6):721-728.

Segurado R, Detera-Wadleigh SD, Levinson DF, Lewis CM, Gill M, Nurnberger JI, Jr., Craddock N, DePaulo JR, Baron M, Gershon ES, Ekholm J, Cichon S, Turecki G, Claes S, Kelsoe JR, Schofield PR, Badenhop RF, Morissette J, Coon H, Blackwood D, McInnes LA, Foroud T, Edenberg HJ, Reich T, Rice JP, Goate A, McInnis MG, McMahon FJ, Badner JA, Goldin LR, Bennett P, Willour VL, Zandi PP, Liu J, Gilliam C, Juo SH, Berrettini WH, Yoshikawa T, Peltonen L, Lonnqvist J, Nothen MM, Schumacher J, Windemuth C, Rietschel M, Propping P, Maier W, Alda M, Grof P, Rouleau GA, Del-Favero J, Van Broeckhoven C, Mendlewicz J, Adolfsson R, Spence MA, Luebbert H, Adams LJ, Donald JA, Mitchell PB, Barden N, Shink E, Byerley W, Muir W, Visscher PM, Macgregor S, Gurling H, Kalsi G, McQuillin A, Escamilla MA, Reus VI, Leon P, Freimer NB, Ewald H, Kruse TA,

Mors O, Radhakrishna U, Blouin JL, Antonarakis SE, Akarsu N. 2003. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. Am J Hum Genet 73(1):49-62.

Simon V, Czobor P, Balint S, Meszaros A, Bitter I. 2009. Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis. Br J Psychiatry 194(3):204-211.

Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N, Edenberg HJ, Nurnberger JI, Rietschel M, Blackwood D, Corvin A, Flickinger M, Guan WH, Mattingsdal M, McQuillin A, Kwan P, Wienker TF, Daly M, Dudbridge F, Holmans PA, Lin DY, Burmeister M, Greenwood TA, Hamshere ML, Muglia P, Smith EN, Zandi PP, Nievergelt CM, McKinney R, Shilling PD, Schork NJ, Bloss CS, Foroud T, Koller DL, Gershon ES, Liu CY, Badner JA, Scheftner WA, Lawson WB, Nwulia EA, Hipolito M, Coryell W, Rice J, Byerley W, McMahon FJ, Schulze TG, Berrettini W, Lohoff FW, Potash JB, Mahon PB, McInnis MG, Zollner S, Zhang P, Craig DW, Szelinger S, Barrett TB, Breuer R, Meier S, Strohmaier J, Witt SH, Tozzi F, Farmer A, McGuffin P, Strauss J, Xu W, Kennedy JL, Vincent JB, Matthews K, Day R, Ferreira MA, O'Dushlaine C, Perlis R, Raychaudhuri S, Ruderfer D, Hyoun PL, Smoller JW, Li J, Absher D, Thompson RC, Meng FG, Schatzberg AF, Bunney WE, Barchas JD, Jones EG, Watson SJ, Myers RM, Akil H, Boehnke M, Chambert K, Moran J, Scolnick E, Djurovic S, Melle I, Morken G, Gill M, Morris D, Quinn E, Muhleisen TW, Degenhardt FA, Mattheisen M, Schumacher J, Maier W, Steffens M, Propping P, Nothen MM, Anjorin A, Bass N, Gurling H, Kandaswamy R, Lawrence J, McGhee K, McIntosh A, McLean AW, Muir WJ, Pickard BS, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Jones IR, Kirov G, Moskvina V, Nikolov I, O'Donovan MC, Owen MJ, Collier DA, Elkin A, Williamson R, Young AH, Ferrier IN, Stefansson K, Stefansson H, Porgeirsson P, Steinberg S, Gustafsson O, Bergen SE, Nimgaonkar V, Hultman C, Landen M, Lichtenstein P, Sullivan P, Schalling M, Osby U, Backlund L, Frisen L, Langstrom N, Jamain S, Leboyer M, Etain B, Bellivier F, Petursson H, Sigurosson E, Muller-Mysok B, Lucae S, Schwarz M, Schofield PR, Martin N, Montgomery GW, Lathrop M, Oskarsson H, Bauer M, Wright A, Mitchell PB, Hautzinger M, Reif A, Kelsoe JR, Purcell SM, D PGCB. 2011. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 43(10):977-U162.

Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI, Ripke S, Santangelo S, Sullivan PF. 2013. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet 381(9875):1371-1379.

Smoller JW, Finn CT. 2003. Family, twin, and adoption studies of bipolar disorder. Am J Med Genet C Semin Med Genet 123C(1):48-58.

Stewart SE, Yu D, Scharf JM, Neale BM, Fagerness JA, Mathews CA, Arnold PD, Evans PD, Gamazon ER, Davis LK, Osiecki L, McGrath L, Haddad S, Crane J, Hezel D, Illman C, Mayerfeld C, Konkashbaev A, Liu C, Pluzhnikov A, Tikhomirov A, Edlund CK, Rauch SL, Moessner R, Falkai P, Maier W, Ruhrmann S, Grabe HJ, Lennertz L, Wagner M, Bellodi L, Cavallini MC, Richter MA, Cook EH, Jr., Kennedy JL, Rosenberg D, Stein DJ, Hemmings SM, Lochner C, Azzam A, Chavira DA, Fournier E, Garrido H, Sheppard B, Umana P, Murphy DL, Wendland JR, Veenstra-VanderWeele J, Denys D, Blom R, Deforce D, Van Nieuwerburgh F, Westenberg HG, Walitza S, Egberts K, Renner T, Miguel EC, Cappi C, Hounie AG, Conceicao do Rosario M, Sampaio AS, Vallada H, Nicolini H, Lanzagorta N, Camarena B, Delorme R, Leboyer M, Pato CN, Pato MT, Voyiaziakis E, Heutink P, Cath DC, Posthuma D, Smit JH, Samuels J, Bienvenu OJ, Cullen B, Fyer AJ, Grados MA, Greenberg BD, McCracken JT, Riddle MA, Wang Y, Coric V, Leckman JF, Bloch M, Pittenger C, Eapen V, Black DW, Ophoff RA, Strengman E, Cusi D, Turiel M, Frau F, Macciardi F, Gibbs JR, Cookson MR, Singleton A, Hardy J, Crenshaw AT, Parkin MA, Mirel DB, Conti DV, Purcell S, Nestadt G, Hanna GL, Jenike MA, Knowles JA, Cox N, Pauls DL. 2013. Genome-wide association study of obsessive-compulsive disorder. Mol Psychiatry 18(7):788-798.

Strother E, Lemberg R, Stanford SC, Turberville D. 2012. Eating disorders in men: underdiagnosed, undertreated, and misunderstood. Eat Disord 20(5):346-355.
Sullivan PF. 1995. Mortality in anorexia nervosa. Am J Psychiatry 152(7):1073-1074.
Sullivan PF, Daly MJ, O'Donovan M. 2012. Genetic architectures of psychiatric disorders: the emerging picture and its implications. Nat Rev Genet 13(8):537-551.
Sullivan PF, Kendler KS, Neale MC. 2003. Schizophrenia as a complex trait - Evidence from a meta-analysis of twin studies. Arch Gen Psychiat 60(12):1187-1192.
Sullivan PF, Neale MC, Kendler KS. 2000. Genetic epidemiology of major depression: Review and meta-analysis. Am J Psychiat 157(10):1552-1562.
Svensson AC, Lichtenstein P, Sandin S, Hultman CM. 2007. Fertility of first-degree relatives of patients with schizophrenia: A three generation perspective. Schizophr Res 91(1-3):238-245.
Szatkiewicz JP, O'Dushlaine C, Chen G, Chambert K, Moran JL, Neale BM, Fromer M, Ruderfer D, Akterin S, Bergen SE, Kahler A, Magnusson PK, Kim Y, Crowley JJ, Rees E, Kirov G, O'Donovan MC, Owen MJ, Walters J, Scolnick E, Sklar P, Purcell S, Hultman CM, McCarroll SA, Sullivan PF. 2014. Copy number variation in schizophrenia in Sweden. Mol Psychiatry.
Thapar A, Collishaw S, Pine DS, Thapar AK. 2012a. Depression in adolescence. Lancet 379(9820):1056-1067.
Thapar A, Cooper M, Jefferies R, Stergiakouli E. 2012b. What causes attention deficit hyperactivity disorder? Arch Dis Child 97(3):260-265.
Thomsen AF, Olsbjerg M, Andersen PK, Kessing LV. 2013. Cohabitation patterns among patients with severe psychiatric disorders in the entire Danish population. Psychol Med 43(5):1013-1021.
Thorgeirsson TE, Gudbjartsson DF, Surakka I, Vink JM, Amin N, Geller F, Sulem P, Rafnar T, Esko T, Walter S, Gieger C, Rawal R, Mangino M, Prokopenko I, Magi R, Keskitalo K, Gudjonsdottir IH, Gretarsdottir S, Stefansson H, Thompson JR, Aulchenko YS, Nelis M, Aben KK, den Heijer M, Dirksen A, Ashraf H, Soranzo N, Valdes AM, Steves C, Uitterlinden AG, Hofman A, Tonjes A, Kovacs P, Hottenga JJ, Willemsen G, Vogelzangs N, Doring A, Dahmen N, Nitz B, Pergadia ML, Saez B, De Diego V, Lezcano V, Garcia-Prats MD, Ripatti S, Perola M, Kettunen J, Hartikainen AL, Pouta A, Laitinen J, Isohanni M, Huei-Yi S, Allen M, Krestyaninova M, Hall AS, Jones GT, van Rij AM, Mueller T, Dieplinger B, Haltmayer M, Jonsson S, Matthiasson SE, Oskarsson H, Tyrfingsson T, Kiemeney LA, Mayordomo JI, Lindholt JS, Pedersen JH, Franklin WA, Wolf H, Montgomery GW, Heath AC, Martin NG, Madden PA, Giegling I, Rujescu D, Jarvelin MR, Salomaa V, Stumvoll M, Spector TD, Wichmann HE, Metspalu A, Samani NJ, Penninx BW, Oostra BA, Boomsma DI, Tiemeier H, van Duijn CM, Kaprio J, Gulcher JR, McCarthy MI, Peltonen L, Thorsteinsdottir U, Stefansson K. 2010. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. Nat Genet 42(5):448-453.
Tielbeek JJ, Medland SE, Benyamin B, Byrne EM, Heath AC, Madden PA, Martin NG, Wray NR, Verweij KJ. 2012. Unraveling the genetic etiology of adult antisocial behavior: a genome-wide association study. Plos One 7(10):e45086.
Trzaskowski M, Eley TC, Davis OS, Doherty SJ, Hanscombe KB, Meaburn EL, Haworth CM, Price T, Plomin R. 2013. First genome-wide association study on anxiety-related behaviours in childhood. Plos One 8(4):e58676.
Uher R. 2009. The role of genetic variation in the causation of mental illness: an evolution-informed framework. Mol Psychiatry 14(12):1072-1082.
Verweij KJ, Vinkhuyzen AA, Benyamin B, Lynskey MT, Quaye L, Agrawal A, Gordon SD, Montgomery GW, Madden PA, Heath AC, Spector TD, Martin NG, Medland SE. 2013. The genetic aetiology of cannabis use initiation: a meta-analysis of genome-wide association studies and a SNP-based heritability estimation. Addict Biol 18(5):846-850.

Vorstman JA, Staal WG, van Daalen E, van Engeland H, Hochstenbach PF, Franke L. 2006. Identification of novel autism candidate regions through analysis of reported cytogenetic abnormalities associated with autism. Mol Psychiatry 11(1):1, 18-28.

Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Bahalim AN, Barker-Collo S, Barrero LH, Bartels DH, Basanez MG, Baxter A, Bell ML, Benjamin EJ, Bennett D, Bernabe E, Bhalla K, Bhandari B, Bikbov B, Bin Abdulhak A, Birbeck G, Black JA, Blencowe H, Blore JD, Blyth F, Bolliger I, Bonaventure A, Boufous S, Bourne R, Boussinesq M, Braithwaite T, Brayne C, Bridgett L, Brooker S, Brooks P, Brugha TS, Bryan-Hancock C, Bucello C, Buchbinder R, Buckle G, Budke CM, Burch M, Burney P, Burstein R, Calabria B, Campbell B, Canter CE, Carabin H, Carapetis J, Carmona L, Cella C, Charlson F, Chen H, Cheng AT, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahiya M, Dahodwala N, Damsere-Derry J, Danaei G, Davis A, De Leo D, Degenhardt L, Dellavalle R, Delossantos A, Denenberg J, Derrett S, Des Jarlais DC, Dharmaratne SD, Dherani M, Diaz-Torne C, Dolk H, Dorsey ER, Driscoll T, Duber H, Ebel B, Edmond K, Elbaz A, Ali SE, Erskine H, Erwin PJ, Espindola P, Ewoigbokhan SE, Farzadfar F, Feigin V, Felson DT, Ferrari A, Ferri CP, Fevre EM, Finucane MM, Flaxman S, Flood L, Foreman K, Forouzanfar MH, Fowkes FG, Franklin R, Fransen M, Freeman MK, Gabbe BJ, Gabriel SE, Gakidou E, Ganatra HA, Garcia B, Gaspari F, Gillum RF, Gmel G, Gosselin R, Grainger R, Groeger J, Guillemin F, Gunnell D, Gupta R, Haagsma J, Hagan H, Halasa YA, Hall W, Haring D, Haro JM, Harrison JE, Havmoeller R, Hay RJ, Higashi H, Hill C, Hoen B, Hoffman H, Hotez PJ, Hoy D, Huang JJ, Ibeanusi SE, Jacobsen KH, James SL, Jarvis D, Jasrasaria R, Jayaraman S, Johns N, Jonas JB, Karthikeyan G, Kassebaum N, Kawakami N, Keren A, Khoo JP, King CH, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lalloo R, Laslett LL, Lathlean T, Leasher JL, Lee YY, Leigh J, Lim SS, Limb E, Lin JK, Lipnick M, Lipshultz SE, Liu W, Loane M, Ohno SL, Lyons R, Ma J, Mabweijano J, MacIntyre MF, Malekzadeh R, Mallinger L, Manivannan S, Marcenes W, March L, Margolis DJ, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGill N, McGrath J, Medina-Mora ME, Meltzer M, Mensah GA, Merriman TR, Meyer AC, Miglioli V, Miller M, Miller TR, Mitchell PB, Mocumbi AO, Moffitt TE, Mokdad AA, Monasta L, Montico M, Moradi-Lakeh M, Moran A, Morawska L, Mori R, Murdoch ME, Mwaniki MK, Naidoo K, Nair MN, Naldi L, Narayan KM, Nelson PK, Nelson RG, Nevitt MC, Newton CR, Nolte S, Norman P, Norman R, O'Donnell M, O'Hanlon S, Olives C, Omer SB, Ortblad K, Osborne R, Ozgediz D, Page A, Pahari B, Pandian JD, Rivero AP, Patten SB, Pearce N, Padilla RP, Perez-Ruiz F, Perico N, Pesudovs K, Phillips D, Phillips MR, Pierce K, Pion S, Polanczyk GV, Polinder S, Pope CA, 3rd, Popova S, Porrini E, Pourmalek F, Prince M, Pullan RL, Ramaiah KD, Ranganathan D, Razavi H, Regan M, Rehm JT, Rein DB, Remuzzi G, Richardson K, Rivara FP, Roberts T, Robinson C, De Leon FR, Ronfani L, Room R, Rosenfeld LC, Rushton L, Sacco RL, Saha S, Sampson U, Sanchez-Riera L, Sanman E, Schwebel DC, Scott JG, Segui-Gomez M, Shahraz S, Shepard DS, Shin H, Shivakoti R, Singh D, Singh GM, Singh JA, Singleton J, Sleet DA, Sliwa K, Smith E, Smith JL, Stapelberg NJ, Steer A, Steiner T, Stolk WA, Stovner LJ, Sudfeld C, Syed S, Tamburlini G, Tavakkoli M, Taylor HR, Taylor JA, Taylor WJ, Thomas B, Thomson WM, Thurston GD, Tleyjeh IM, Tonelli M, Towbin JA, Truelsen T, Tsilimbaris MK, Ubeda C, Undurraga EA, van der Werf MJ, van Os J, Vavilala MS, Venketasubramanian N, Wang M, Wang W, Watt K, Weatherall DJ, Weinstock MA, Weintraub R, Weisskopf MG, Weissman MM, White RA, Whiteford H, Wiersma ST, Wilkinson JD, Williams HC, Williams SR, Witt E, Wolfe F, Woolf AD, Wulf S, Yeh PH, Zaidi AK, Zheng ZJ, Zonies D, Lopez AD, Murray CJ, AlMazroa MA, Memish ZA. 2012. Years lived with disability (YLDs) for

1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380(9859):2163-2196.

Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, Nord AS, Kusenda M, Malhotra D, Bhandari A, Stray SM, Rippey CF, Roccanova P, Makarov V, Lakshmi B, Findling RL, Sikich L, Stromberg T, Merriman B, Gogtay N, Butler P, Eckstrand K, Noory L, Gochman P, Long R, Chen ZG, Davis S, Baker C, Eichler EE, Meltzer PS, Nelson SF, Singleton AB, Lee MK, Rapoport JL, King MC, Sebat J. 2008. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 320(5875):539-543.

Wang K, Zhang H, Bloss CS, Duvvuri V, Kaye W, Schork NJ, Berrettini W, Hakonarson H. 2011. A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. Mol Psychiatry 16(9):949-959.

Williams NM, Franke B, Mick E, Anney RJ, Freitag CM, Gill M, Thapar A, O'Donovan MC, Owen MJ, Holmans P, Kent L, Middleton F, Zhang-James Y, Liu L, Meyer J, Nguyen TT, Romanos J, Romanos M, Seitz C, Renner TJ, Walitza S, Warnke A, Palmason H, Buitelaar J, Rommelse N, Vasquez AA, Hawi Z, Langley K, Sergeant J, Steinhausen HC, Roeyers H, Biederman J, Zaharieva I, Hakonarson H, Elia J, Lionel AC, Crosbie J, Marshall CR, Schachar R, Scherer SW, Todorov A, Smalley SL, Loo S, Nelson S, Shtir C, Asherson P, Reif A, Lesch KP, Faraone SV. 2012. Genome-wide analysis of copy number variants in attention deficit hyperactivity disorder: the role of rare variants and duplications at 15q13.3. Am J Psychiatry 169(2):195-204.

Williams NM, Zaharieva I, Martin A, Langley K, Mantripragada K, Fossdal R, Stefansson H, Stefansson K, Magnusson P, Gudmundsson OO, Gustafsson O, Holmans P, Owen MJ, O'Donovan M, Thapar A. 2010. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. Lancet 376(9750):1401-1408.

Woodley MA. 2009. Inbreeding depression and IQ in a study of 72 countries. Intelligence 37(3):268-276.

Wray NR, Gottesman, II. 2012. Using summary data from the danish national registers to estimate heritabilities for schizophrenia, bipolar disorder, and major depressive disorder. Frontiers in genetics 3:118.

Xie P, Kranzler HR, Yang C, Zhao H, Farrer LA, Gelernter J. 2013. Genome-wide association study identifies new susceptibility loci for posttraumatic stress disorder. Biol Psychiatry 74(9):656-663.

Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. 2008. Strong association of de novo copy number mutations with sporadic schizophrenia. Nat Genet 40(7):880-885.

Yang JA, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: A Tool for Genome-wide Complex Trait Analysis. Am J Hum Genet 88(1):76-82.

Yang L, Neale BM, Liu L, Lee SH, Wray NR, Ji N, Li H, Qian Q, Wang D, Li J, Faraone SV, Wang Y, Doyle AE, Reif A, Rothenberger A, Franke B, Sonuga-Barke EJ, Steinhausen HC, Buitelaar JK, Kuntsi J, Biederman J, Lesch KP, Kent L, Asherson P, Oades RD, Loo SK, Nelson SF, Smalley SL, Banaschewski T, Arias Vasquez A, Todorov A, Charach A, Miranda A, Warnke A, Thapar A, Cormand B, Freitag C, Mick E, Mulas F, Middleton F, HakonarsonHakonarson H, Palmason H, Schafer H, Roeyers H, McGough JJ, Romanos J, Crosbie J, Meyer J, Ramos-Quiroga JA, Sergeant J, Elia J, Langely K, Nisenbaum L, Romanos M, Daly MJ, Ribases M, Gill M, O'Donovan M, Owen M, Casas M, Bayes M, Lambregts-Rommelse N, Williams N, Holmans P, Anney RJ, Ebstein RP, Schachar R, Medland SE, Ripke S, Walitza S, Nguyen TT, Renner TJ, Hu X. 2013. Polygenic transmission and complex neuro developmental network for attention deficit hyperactivity disorder: genome-wide association study of both common and rare variants. Am J Med Genet B Neuropsychiatr Genet 162B(5):419-430.

Zammit S, Allebeck P, Dalman C, Lundberg I, Hemmingson T, Owen MJ, Lewis G. 2003. Paternal age and risk for schizophrenia. Br J Psychiatry 183:405-408.

Zhang D, Cheng L, Qian Y, Alliey-Rodriguez N, Kelsoe JR, Greenwood T, Nievergelt C, Barrett TB, McKinney R, Schork N, Smith EN, Bloss C, Nurnberger J, Edenberg HJ, Foroud T, Sheftner W, Lawson WB, Nwulia EA, Hipolito M, Coryell W, Rice J, Byerley W, McMahon F, Schulze TG, Berrettini W, Potash JB, Belmonte PL, Zandi PP, McInnis MG, Zollner S, Craig D, Szelinger S, Koller D, Christian SL, Liu C, Gershon ES. 2009. Singleton deletions throughout the genome increase risk of bipolar disorder. Mol Psychiatry 14(4):376-380.

Zietsch BP, Verweij KJ, Heath AC, Martin NG. 2011. Variation in human mate choice: simultaneously investigating heritability, parental influence, sexual imprinting, and assortative mating. Am Nat 177(5):605-616.

Zipfel S, Lowe B, Reas DL, Deter HC, Herzog W. 2000. Long-term prognosis in anorexia nervosa: lessons from a 21-year follow-up study. Lancet 355(9205):721-722.